**CARINE RODRIGUES PEREIRA**

# GENOMIC EPIDEMIOLOGY OF *Brucella abortus* ISOLATED FROM CATTLE IN BRAZIL

**LAVRAS – MG**
**2023**

**CARINE RODRIGUES PEREIRA**

# GENOMIC EPIDEMIOLOGY OF *Brucella abortus* ISOLATED FROM CATTLE IN BRAZIL

Thesis presented to the Universidade Federal de Lavras, as part of the requirements of the Graduate Program in Veterinary Sciences, to obtain the title of PhD.

Dra. Elaine Maria Seles Dorneles
Supervisor

Dr. Vasco Ariston de Carvalho Azevedo
Co supervisor

**LAVRAS – MG**
**2023**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Pereira, Carine Rodrigues.
Genomic epidemiology of *Brucella abortus* isolated from cattle in Brazil / Carine Rodrigues Pereira. - 2023.
105 p.: il.

Orientador(a): Elaine Maria Seles Dorneles.
Coorientador(a): Vasco Ariston de Carvalho Azevedo.
Tese (doutorado) - Universidade Federal de Lavras, 2023.
Bibliografia.

1. Brucelose. 2. Sequenciamento completo do genoma. 3. Vigilância. I. Dorneles, Elaine Maria Seles. II. Azevedo, Vasco Ariston de Carvalho. III. Título.

**CARINE RODRIGUES PEREIRA**

# GENOMIC EPIDEMIOLOGY OF *Brucella abortus* ISOLATED FROM CATTLE IN BRAZIL

# EPIDEMIOLOGIA GENÔMICA DE *Brucella abortus* ISOLADAS DE BOVINOS NO BRASIL

Thesis presented to the Universidade Federal de Lavras, as part of the requirements of the Graduate Program in Veterinary Sciences, to obtain the title of PhD.

Approved on June 23, 2023.

| | |
|---|---|
| Dra. Elaine Maria Seles Dorneles | UFLA |
| Dr. David O'Callaghan | UM |
| Dr. Andrey Pereira Lage | UFMG |
| Dr. Marcus Vinicius Canário Viana | UFMG |
| Dr. Siomar de Castro Soares | UFTM |

gov.br
Documento assinado digitalmente
ELAINE MARIA SELES DORNELES
Data: 29/07/2023 07:22:30-0300
Verifique em https://validar.iti.gov.br

Dra. Elaine Maria Seles Dorneles
Supervisor

**LAVRAS – MG**
**2023**

Dedico às minhas tias Hilda (*in memoriam*) e Neusa, que sempre me incentivaram a perseverar nos estudos e são meus exemplos na busca pelo conhecimento.

## AGRADECIMENTOS

A Deus por todas as bênçãos durante esta jornada.

A toda minha família, em especial aos meus pais Valéria e Geraldo, avós e tias-avós por sempre me incentivarem para seguir o meu sonho de obter este diploma.

À Elaine por toda orientação, ensinamentos, carinho e amizade.

Aos professores Andrey e Vasco pela generosidade e oportunidades concedidas.

To Jeff and David for their willingness to contribute to my formation and expand my horizons.

A toda a equipe RECOM, em especial Alessandra, Andrei, Rose, professor Siomar e Thiago por todo treinamento, paciência e colaboração.

À Raquel, Saulo e Laís por toda ajuda e empenho, na orientação de vocês aprendi muito mais que ensinei.

À Camila e Jonatã por todo suporte e amizade durante meu período nos EUA.

À Dani, por todo apoio e direcionamento durante toda minha história na UFLA.

To all my friends in Flagstaff especially Li, Irene, Erik and Benjamin for making me feel so welcome during my stay.

Ao Lucas e Raíssa por sempre se fazerem presentes mesmo com toda distância.

Aos amigos do LEM, em especial Anna, Dirce, Maysa e Rafa pela convivência diária (ainda que online) e por tornarem esta etapa ainda mais especial.

À UFLA, especialmente todos os funcionários do DMV por serem tão cooperativos em me ajudar a atingir meu objetivo.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Fundação de Amparo à Pesquisa do Estado de Minas Gerais pelo apoio financeiro.

## ABSTRACT

Bovine brucellosis is a bacterial zoonosis that affects livestock and public health. In Brazil, it is caused mainly by *Brucella abortus*. The disease causes reproductive clinical signs in cattle, such as abortions and premature births, and nonspecific clinical signs in humans, as fever, arthralgia, night sweats, among others. In humans, the treatment is difficult, and the combination of synergistic antimicrobials for periods of 4 to 6 weeks is preconized. The aims of this study were (1) to perform a systematic review of antimicrobial resistance (AMR) in *Brucella* spp.; (2) to identify the genetic mechanisms related to AMR in *Brucella abortus* isolated from cattle in Brazil; (3) to compare the population structure of the genomes of 53 Brazilian *B. abortus* isolates using eight different genotyping methods; (4) to perform a pan-genome analysis of this species and (5) to describe the identification and complete sequencing of the first strain of *Pseudochrobactrum saccharolyticum* isolated in Latin America, previously classified as *B. abortus*. It was possible to identify that *Brucella* spp. is mainly resistant to rifampicin and aminoglycosides, the mechanisms of genetic resistance in the genus are still poorly understood, and this could be attributed to the high number of hypothetical proteins that remain with their unknown function in this pathogen. In addition, it is essential that the solutions regarding any questions using the implementation of bioinformatics methods always consider the epidemiological context of a isolated strain. Genomics was a fundamental tool to answer several questions related to *B. abortus* strains from Brazil and enabled the identification of a genus from *Brucellaceae* family hitherto never found in the country. The intercession between the epidemiological and whole genome sequence information of the strains investigated provided important information on specific nuances of the *Brucellaceae* family, especially *B. abortus*, providing a better understanding of this pathogen of worldwide importance.



**Key-words:** Brucellosis. Whole genome sequencing. Surveillance. Antimicrobial resistance

## RESUMO

A brucelose bovina é uma zoonose bacteriana que afeta a pecuária e a saúde pública. No Brasil, é causada principalmente pela *Brucella abortus*. A doença causa sinais clínicos reprodutivos em bovinos, como abortos e partos prematuros, e sinais clínicos inespecíficos em humanos, como febre, artralgia, sudorese noturna, entre outros. Em humanos, o tratamento é difícil, preconizando-se a combinação de antimicrobianos sinérgicos por períodos de 4 a 6 semanas. Os objetivos deste estudo foram (1) realizar uma revisão sistemática da resistência antimicrobiana (AMR) em *Brucella* spp.; (2) identificar os mecanismos genéticos relacionados à AMR em *Brucella abortus* isolados de bovinos no Brasil; (3) comparar a estrutura populacional dos genomas de 53 isolados brasileiros de *B. abortus* usando oito diferentes métodos de genotipagem; (4) realizar uma análise pan-genômica desta espécie e (5) descrever a identificação e sequenciamento completo da primeira cepa de *Pseudochrobactrum saccharolyticum* isolada na América Latina, previamente classificada como *B. abortus*. Foi possível identificar que *Brucella* spp. é principalmente resistente à rifampicina e aos aminoglicosídeos, os mecanismos de resistência genética no gênero ainda são pouco compreendidos, e isso pode ser atribuído ao elevado número de proteínas hipotéticas que permanecem com função desconhecida neste patógeno. Além disso, é fundamental que as soluções de quaisquer dúvidas com a aplicação de métodos de bioinformática sempre considerem o contexto epidemiológico de uma cepa isolada. A genômica foi uma ferramenta fundamental para responder a diversas questões relacionadas às cepas de *B. abortus* do Brasil e possibilitou a identificação de um gênero da família *Brucellaceae* até então nunca encontrado no país. A intercessão entre as informações epidemiológicas e do sequenciamento completo do genoma das cepas investigadas forneceu informações importantes sobre nuances específicas da família *Brucellaceae*, especialmente *B. abortus*, proporcionando um melhor entendimento deste patógeno de importância mundial.



**Palavras-chave:** Brucelose. Sequenciamento completo do genoma. Vigilância. Resistência antimicrobiana

**SUMMARY**

## GENERAL INTRODUCTION

Brucellosis is one of the most common anthropozoonosis in the world, with approximately 500,000 new human cases reported annually to the World Health Organization (WHO), besides being one of the main causes of economic losses in livestock (KIIZA; DENAGAMAGE; SERRA; MAUNSELL *et al.*, 2023; PAPPAS; PAPADIMITRIOU; AKRITIDIS; CHRISTOU *et al.*, 2006). Thirteen species are described within the genus *Brucella*, being *Brucella abortus* the main responsible for causing the disease in cattle (ABOUT; PASTRE; BOUTROU; MARTINEZ *et al.*, 2023; WHATMORE; FOSTER; EVOLUTION, 2021). In Brazil, bovine brucellosis is endemic and present in all states, with the prevalence of positive properties ranging from 0.91% [Confidence interval (CI) 95% (0.30 - 2.11)] in Santa Catarina (BAUMGARTEN; VELOSO; GRISI-FILHO; FERREIRA *et al.*, 2016) to 30.6% [CI 95% (27.4 - 34.0)] in Mato Grosso do Sul (LEAL FILHO; BOTTENE; MONTEIRO; PELLEGRIN *et al.*, 2016). In 2001, the National Program for the Control and Eradication of Brucellosis and Animal Tuberculosis (PNCEBT) was created, aiming to mitigate the considerable economic losses caused by brucellosis and to reduce the transmission of the disease to humans (BRASIL, 2001). The program is based on the following strategies: compulsory brucellosis vaccination of heifers, diagnosis of animals in transit, slaughtering of positive animals, and certification of disease free properties (BRASIL, 2017).

Although methodologies that include complete genome sequencing have been successfully implemented to better characterize *Brucella* spp. isolated in several countries (GEORGI; WALTER; PFALZGRAF; NORTHOFF *et al.*, 2017; LI; WANG; ZHU; WANG *et al.*, 2020; SUÁREZ-ESQUIVEL; HERNÁNDEZ-MORA; RUIZ-VILLALOBOS; BARQUERO-CALVO *et al.*, 2020), this approach is still not widely implemented to deal with zoonotic diseases, such as brucellosis, in Brazil. The identification, sequencing, and genetic characterization of *B. abortus* are necessary strategies to identify the source of infection and to plan correct control and prevention measures. Recent advances in sequencing technologies and bioinformatics tools have made this methodology a viable and innovative solution for epidemiological investigation and surveillance of pathogens (KAO; HAYDON; LYCETT; MURCIA, 2014).

The term genomic epidemiology has been increasingly used to describe the use of complete bacterial genome sequencing to access and analyze the DNA sequence resources with epidemiological importance. Genomic elements and single nucleotide polymorphisms (SNPs) in bacterial evolution are important targets for epidemiological investigations at different time

and geographic scales (DENG; BAKKER; HENDRIKSEN, 2016). The whole genome sequencing of *B. abortus* strains opens several possibilities for investigations with great potential to guide the control and prevention measures used by PNCEBT in Brazil. The association of this approach with bioinformatics tools allows the organization and interpretation of these large data sets, generating a great potential for the application of these data in epidemiological problems that requires high-resolution. Information obtained using these tools can be applied very effectively to improve the understanding of the mechanisms and patterns related to evolutionary processes, acquisition of resistance to antimicrobials, and transmission dynamics observed in *B. abortus* isolated in Brazil.

## REFERENCES


ABOUT, F.; PASTRE, T.; BOUTROU, M.; MARTINEZ, A. Y. *et al.* Novel Species of *Brucella* Causing Human Brucellosis, French Guiana. **Emerg Infect Dis**, 29, n. 2, p. 333-340, Feb 2023.

BAUMGARTEN, K. D.; VELOSO, F. P.; GRISI-FILHO, J. H. H.; FERREIRA, F. *et al.* Prevalence and risk factors for bovine brucellosis in the State of Santa Catarina, Brazil. **Semina: Ciências Agrárias**, 37, n. 5Supl2, p. 3425, 2016.

BRASIL. Instrução Normativa n. 2 de 10 de janeiro de 2001. MINISTÉRIO DA AGRICULTURA, P. e. A. Brasília, Distrito Federal: Diário Oficial da União**:** 26-31 p. 2001.

BRASIL. Instrução normativa nº 10, de 3 de março de 2017. MINISTÉRIO DA AGRICULTURA, P. e. A. Brasília, Distrito Federal: Diário Oficial da União**:** 4-8 p. 2017.

DENG, X.; BAKKER, H. C. d.; HENDRIKSEN, R. S. Genomic Epidemiology: Whole-Genome-Sequencing–Powered Surveillance and Outbreak Investigation of Foodborne Bacterial Pathogens. **Annual Review of Food Science and Technology**, 7, n. 1, p. 353-374, 2016.

GEORGI, E.; WALTER, M. C.; PFALZGRAF, M. T.; NORTHOFF, B. H. *et al.* Whole genome sequencing of *Brucella melitensis* isolated from 57 patients in Germany reveals high diversity in strains from Middle East. **PLoS One**, 12, n. 4, 2017.

KAO, R. R.; HAYDON, D. T.; LYCETT, S. J.; MURCIA, P. R. Supersize me: how whole-genome sequencing and big data are transforming epidemiology. **Trends in Microbiology**, 22, n. 5, p. 282-291, 2014/05/01/ 2014.

KIIZA, D.; DENAGAMAGE, T.; SERRA, R.; MAUNSELL, F. *et al.* A systematic review of economic assessments for brucellosis control interventions in livestock populations. **Prev Vet Med**, 213, p. 105878, Apr 2023.

LEAL FILHO, J. M.; BOTTENE, I. F. N.; MONTEIRO, L. A. R. C.; PELLEGRIN, A. O. *et al.* Control of bovine brucellosis from 1998 to 2009 in the state of Mato Grosso do Sul, Brazil. **Semina: Ciências Agrárias**, 37, n. 5, p. 3467-3477, 2016.

LI, Z.; WANG, X.-m.; ZHU, X.; WANG, M. *et al.* Molecular Characteristics of *Brucella* Isolates Collected From Humans in Hainan Province, China. **Frontiers in Microbiology**, 11, n. 452, 2020-March-27 2020. Original Research.

PAPPAS, G.; PAPADIMITRIOU, P.; AKRITIDIS, N.; CHRISTOU, L. *et al.* The new global map of human brucellosis. **The Lancet Infectious Diseases**, 6, n. 2, p. 91-99, 2006.

SUÁREZ-ESQUIVEL, M.; HERNÁNDEZ-MORA, G.; RUIZ-VILLALOBOS, N.; BARQUERO-CALVO, E. *et al.* Persistence of *Brucella abortus* lineages revealed by genomic characterization and phylodynamic analysis. **PLOS Neglected Tropical Diseases**, 14, n. 4, p. e0008235, 2020.

WHATMORE, A. M.; FOSTER, J. T. Emerging diversity and ongoing expansion of the genus *Brucella*. **Infection, Genetics and Evolution**, 92, p. 104865, 2021.

# CHAPTER – 1: A systematic review of antimicrobial resistance in *Brucella* spp.

Carine Rodrigues Pereira[1]; Laís Gabrielle Alvarenga Assis[1]; Jeffrey T. Foster[2]; David O'Callaghan[3]; Andrey Pereira Lage[4]; Elaine Maria Seles Dorneles[1]*

[1]Faculdade de Zootecnia e Medicina Veterinária, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil

[2]Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, Arizona, United States of America

[3]Bacterial Virulence and Infectious Disease, University of Montpellier, Nimes, France

[4]Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

*Address correspondence to author: +55 (35) 3829-1712 – elaine.dorneles@ufla.br – Universidade Federal de Lavras Caixa Postal 3037 – CEP: 37200-900 – Lavras – MG – Brazil

## ABSTRACT

Brucellosis is a zoonotic disease of remarkable importance worldwide. Each year more cases of antimicrobial resistance (AMR) have been reported. The focus of this systematic review was to report the main phenotypic methods and antimicrobials investigated in *Brucella* spp. drug susceptibility tests, as well as to identify the main genetic mechanisms investigated for this outcome. Seven databases were used to identify papers related to AMR in *Brucella* spp.: CABI, Cochrane, Proquest, Pubmed, Scielo, Science Direct, and Web of Science. The search resulted in 3,444 studies, of which 62 were selected based on the exclusion/inclusion and quality criteria established by this systematic review. The main phenotypic method used to assess drug susceptibility among the selected papers was E-test, whereas the most tested drugs were rifampicin, doxycycline, and ciprofloxacin, being rifampicin the one with the highest number of studies that reported resistance. The genes reported as associated with AMR were *aadA*, *rpoB*, *parC*, *gyrB*, *gyrA*, *23s rRNA*, *norM*, *bepR*, *bepC*, *bepG*, *bepE*, *DacC*, *potA*, and *VirB*. This systematic review improved the understanding of AMR in *Brucella* spp.; however, the is a lack of standardization of phenotypic methods for measuring drug susceptibility in *Brucella* spp. imposed a limitation on the summarization of the data and advances in this field. Considering the investigation of genetic markers for AMR in *Brucella* spp., the whole genome approach

appeared to be more effective than sequencing methods limited to genes previously reported to cause resistance in other pathogens.

**KEY WORDS:** Brucellosis, susceptibility, antimicrobials, multidrug resistance, SNPs

## INTRODUCTION

Millions of people are routinely exposed to *Brucella* spp. throughout their everyday lives and contract brucellosis each year worldwide, especially at Latin America, the Middle East, Africa, and Asia (LAINE; SCOTT; ARENAS-GAMBOA, 2022). This zoonotic bacterial disease also affects animals, and results in tremendous economic losses through reproductive failure in livestock (KIIZA; DENAGAMAGE; SERRA; MAUNSELL *et al.*, 2023). Brucellosis in man is usually caused by *B. melitensis*, *B. abortus* or *B. suis* mainly through exposure to *Brucella*-contaminated milk and secretions from infected animals (PAPPAS; AKRITIDIS; BOSILKOVSKI; TSIANOS, 2005). Human brucellosis has a broad spectrum of clinical manifestations, the acute form it usually characterized by febrile illness, fatigue, anorexia, weight loss and generalized aching, whereas in the chronic form it may causes multisystemic severe complications (DEAN; CRUMP; GRETER; HATTENDORF *et al.*, 2012). In fact, the Disability-Adjusted Life Year (DALY), a metric that quantifies the burden of mortality and morbidity caused by a disease, were found to be 0.13 [95% uncertainty interval (UI) 0.06–0.18] per thousand people per year for human brucellosis (SINGH; KHATKAR; AULAKH; GILL *et al.*, 2018).

Since *Brucella* spp. is a facultative intracellular bacterium, few antimicrobial combinations are effective to penetrate macrophages and have a good action against this organism (ALISKAN; CAN; DEMIRBILEK; COLAKOGLU *et al.*, 2009). The World Health Organization (WHO) preconizes the use of dual or triple regimens of drugs as doxycycline, streptomycin, gentamycin, rifampicin, tetracycline and co-trimoxazole for at least six weeks to treat human brucellosis (CORBEL, 2006). Although *Brucella* isolates are generally considered susceptible to these antimicrobials, each year more cases of resistance and disease relapses have been reported (ABDEL-MAKSOUD; HOUSE; WASFY; ABDEL-RAHMAN *et al.*, 2012; ASADI; HASHEMI; YOUSEF ALIKHANI; MOGHIMBEIGI *et al.*, 2017; DESHMUKH; HAGEN; AL SHARABASI; ABRAHAM *et al.*, 2015; HASHIM; AHMAD; MOHAMED ZAHIDI; TAY *et al.*, 2014; JIANG; MAO; ZHAO; LI *et al.*, 2010; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018; KHAN; SHELL; MELZER; SAYOUR *et al.*, 2019; LIU; DI;

WANG; LIU *et al.*, 2018; MAGALHÃES NETO; CORÇÃO; DASSO; KEID *et al.*, 2014; PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015; SHEVTSOV; SYZDYKOV; KUZNETSOV; SHUSTOV *et al.*, 2017; WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021). Considering that *in vitro* antimicrobial susceptibilities can change over time and from one geographic region to another, tests should be performed more frequently to assess its spread in *Brucella* spp. However, susceptibility tests are still poorly standardized for *Brucella* species and are not routinely performed, especially due to biosafety level 3 facilities requirement (BAYKAM; ESENER; ERGÖNÜL; EREN *et al.*, 2004).

As an alternative to phenotypic methods, the genetic approach for identifying determinants involved in the emergence of antimicrobial resistance (AMR) in *Brucella* has been increasingly explored (ALJANAZREH; ALZATARI; TAMIMI; ALSAAFEEN *et al.*, 2021; DADAR; ALAMIAN; BRANGSCH; ELBADAWY *et al.*, 2023; ELBEHIRY; ALDUBAIB; AL RUGAIE; MARZOUK *et al.*, 2022; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018; KHAZAEI; NAJAFI; PIRANFAR; MIRNEJAD, 2016; MARIANELLI; CIUCHINI; TARANTINO; PASQUALI *et al.*, 2004; MARTIN; POSADAS; CARRICA; CRAVERO *et al.*, 2009). Nonetheless, as observed for phenotypic tests, this approach also has some limitations. Indeed, the understanding genetic determinants of resistance for this bacteria very incipient, since its genome is extremely clonal and not prone to horizontal transfer (SUÁREZ-ESQUIVEL, M.; CHAVES-OLARTE, E.; MORENO, E.; GUZMÁN-VERRI, C., 2020), which makes the identification of the mechanisms responsible for this phenotype a real challenge (PEREIRA; KATO; ARAÚJO; DA SILVA *et al.*, 2023).

Since AMR in *Brucella* is an area in full expansion at the frontier of knowledge, a systematic review can elucidate and establish some patterns related to drug resistance detected so far, as well as direct researchers regarding quality standards for the correct measurement of this phenomenon not only in *Brucella* but also in other microorganisms. Thus, this study aimed to investigate which are the main phenotypic and genotypic methods used to determine antimicrobial susceptibility in *Brucella* spp. and which antimicrobials and genes have been investigated at this genus related to AMR.

## MATERIAL AND METHODS

The guidelines of PRISMA statement (Preferred Reported Items for Systematic Reviews and Meta-Analyses) (PAGE; MCKENZIE; BOSSUYT; BOUTRON *et al.*, 2021) were

formally adopted in this review and can be seen in additional file 1 (S1 Appendix). The review was also registered as PROSPERO – International prospective register of systematic reviews (ID: 300883).

**Search strategy**

The search was conducted in December 2021, without any date or country restriction. All the keywords were investigated within title, abstract and full text sections in the following databases: CABI, Cochrane, Proquest, Pubmed, Scielo, Science Direct, and Web of Science (S2 Appendix). Briefly, the PICO (population, intervention, comparison and outcome) involved (i) *Brucella* spp., (ii) antimicrobial/antibiotic used to treat human brucellosis, (iii) susceptibility/sensitivity, and (iv) resistance. An overview of the search terms is shown in additional file 2 (S2 Appendix).

**Selection of the studies**

After the literature search, the database was exported to Endnote X9, checked and cleaned for duplicates (HUPE, 2019). Then, the records were screened by their titles (CRP) and the remained ones were independently evaluated by two reviewers based on their abstract (CRP and EMSD) using Rayyan (OUZZANI; HAMMADY; FEDOROWICZ; ELMAGARMID, 2016). Subsequently, full text of the papers selected based on the abstracts were evaluated by two reviewers (CRP and EMSD) in terms of its relevance and by means of inclusion/exclusion criteria. When these reviewers disagreed in any stage, a meeting was schedule to reach a consensus. Further, the reference lists of review papers were screened to find pertinent studies not identified during the initial search.

**Inclusion and exclusion criteria**

The following characteristics were considered for the inclusion of articles: (i) all years and countries, (ii) articles focusing on *Brucella* spp., (iii) investigating the phenotype or genotype of AMR (iii) written in English, French, Korean, Spanish or Portuguese. Articles aiming on (i) immunology, vaccination, parasite/host relationship, diagnostic and epidemiology or (ii) or whose full text was not available were excluded. Full inclusion and exclusion criteria are shown in additional file 3 (S3 Appendix).

**Quality assessment and data extraction**

Data were extracted from papers by one of the reviewers (CRP) and subsequently checked for accuracy by another reviewer (LGAA). Disagreements regarding data extraction

among reviewers were solved by consensus. Extracted data included: first author, year of strain isolation, country, *Brucella* species, number field and reference strains investigated, host, biological sample, phenotypic methods used [minimal inhibitory concentration (MIC) and disk diffusion] and its details (protocol and organism of reference, quality control strain, temperature and incubation time, and inoculum concentration), tested antimicrobials (concentrations/range, MIC50 and MIC90, and percentage of resistance), investigated genes, mechanism of resistance and antimicrobial classes related to the assessed gene.

Since, in the best of our knowledge, there is no established criteria (checklist) in the literature to assess the quality of articles that perform phenotypic antimicrobial susceptibility tests, herein we implemented some quality standards to select only high-quality papers. Following the methodological guidelines recommended by the Clinical & Laboratory Standards Institute Guidelines (CLSI/NCCS) and European Committee on Antimicrobial Susceptibility Testing (EUCAST), six parameters were defined and sequentially assessed in each study that performed the phenotypic evaluation of AMR: (i) *Brucella* spp. classification method, (ii) quality control strain, (iii) organism of reference, (iv) incubation temperature, (v) incubation time and (vi) inoculum concentration. Studies that did not clearly address these parameters or that provided unsatisfactory information (outside the recommendations of CLSI/NCCLS or EUCAST manual) were considered of low quality and thereby excluded, while the others that met the eligibility criteria were included. No quality criteria was adopted in the screening of the papers that used only genotypic methods to assess AMR; moreover, genotypic data from papers that did not met the quality criteria for the phenotypic data were also extracted.

**Data synthesis**

Data from the studies remaining at the final stage of the selection process were summarized in tables and figures. The R software version 4.1.2 (R, 2023) was used to plot the graphs and to draw up the maps, with aid of the packages ggplot2 and plotly (SIEVERT, 2020; WICKHAM, 2011).

## RESULTS

The search strategy adopted identified a total of 3,444 papers, published between 1949 and 2022; 1,335 duplicates were excluded, and 418 full texts were assessed for eligibility. These reports were evaluated and 208 were excluded with reason. Additionally, 47 studies were retrieved from the reviews' references, being 8 of them included at the quality assessment step.

The 218 studies (210 + 8) included for full text selection were divided in three groups: the ones that evaluated only phenotype (184), only genotype (10) or both outcomes (24). The papers that evaluated AMR phenotype were selected based on the established quality criteria: 156 studies from the 183 that assessed only phenotype and 14 papers that evaluated both outcomes were excluded, the latter were included only for genotyping assessment purposes. Subsequently, 62 papers, from 2002 to 2022, were included in quality level assessment and data synthesis appraisal, after a thorough review (Fig. 1). The background characteristics (geographic location, study period, *Brucella* species, host, number investigated strains, phenotypic method used, genotypic method used, tested antimicrobials, and investigated genes) were extracted from these articles and are shown in the additional file 4 (S4 Appendix). The studies excluded based on the quality parameters established and reasons for the exclusion are listed in the additional file 5 (S5 Appendix).

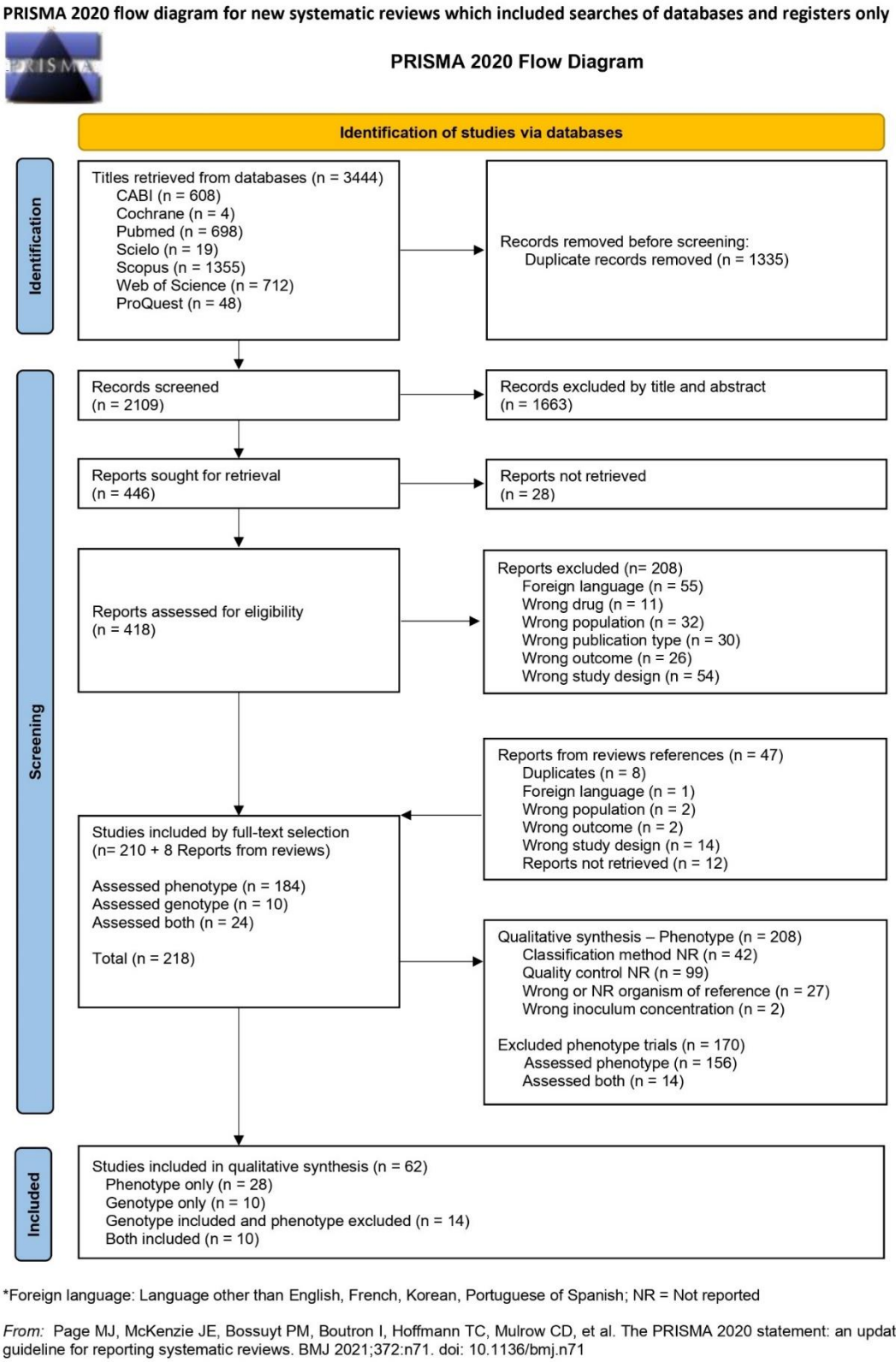


Figure 1: PRISMA Flow diagram of selected studies.

**Study characteristics**

The assessment of geographical origin of the 62 selected papers showed that from the 3,800 assessed field strains (phenotypic and genotypic methods), 2,854 were from Asia, 448

from Africa, 292 from Europe and 206 from America (Fig 2). Blood (3,064), aborted fetuses (75) and synovial fluids (64) were the major samples collected for *Brucella* spp. isolation followed by other (190) and not reported (407) sources, with the most investigated species being *B. melitensis* (2,785), *B. abortus* (216), and *B. suis* (2), in addition to 797 strains identified only at the genus level. Fig. 3 shows the distribution of the analyzed strains according to their host (a), clinical sample (b) and the *Brucella* species classification (c).

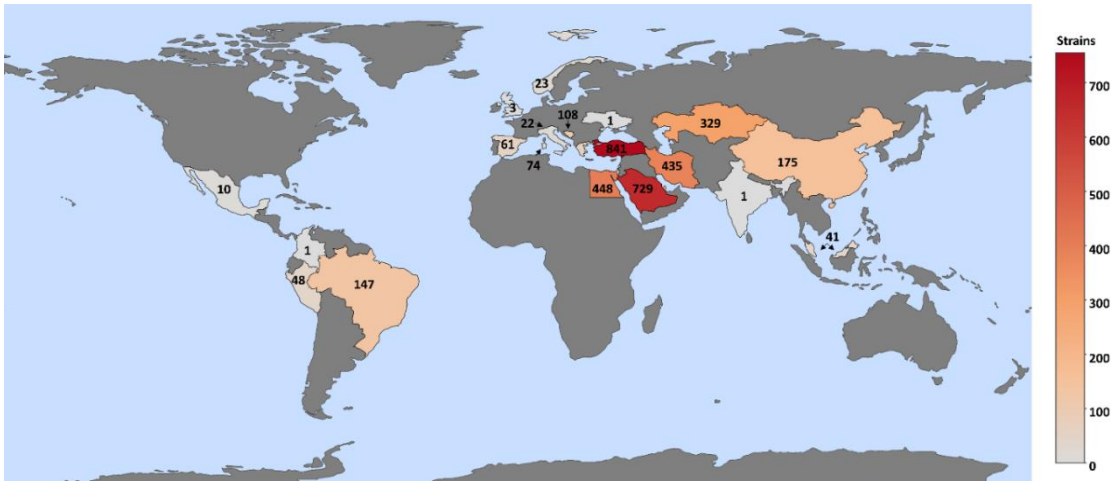


Figure 2: Geographical distribution of the field strains investigated for antimicrobial susceptibility in the papers included in the present review.

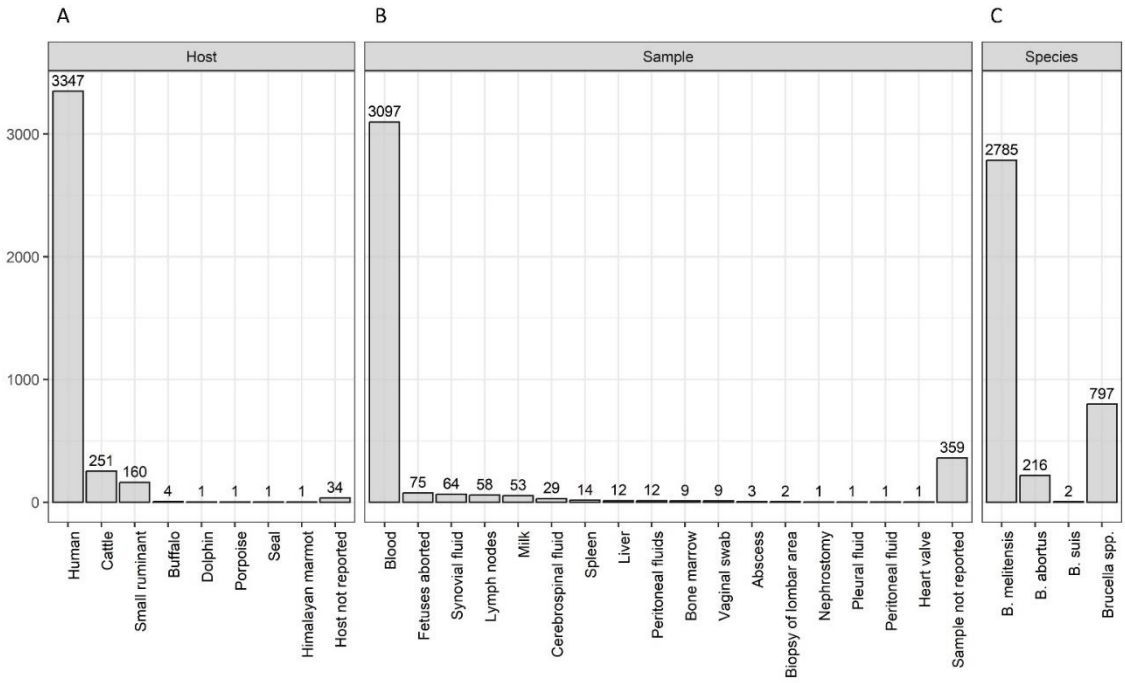


Figure 3: Distribution of the field strains investigated for antimicrobial susceptibility in the papers included in the present review, according to their host (a), clinical sample (b) and the *Brucella* species classification (c).

**Phenotype investigation**

The MIC test, including E-test (28) and dilution tests (10) (micro and microdilution), were the most used methods to assess susceptibility of the *Brucella* spp., followed by the disk diffusion test (4). Of the 38 studies that phenotypically evaluated this outcome attending the quality parameters, 4 used more than one method (Fig. 4A). A total of 34 antimicrobials were tested, being rifampicin, doxycycline, and ciprofloxacin, the most tested drugs in the quantitative method (MIC) (Fig. 4B) (S4 Appendix).

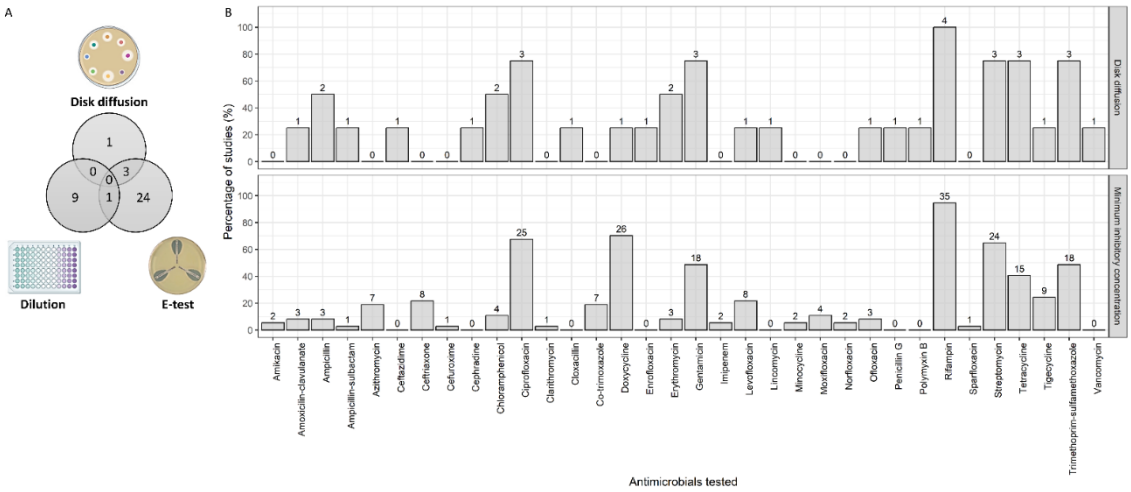


Figure 4: Methods employed (a) and antimicrobials assessed among the papers included in the present review that performed antimicrobial susceptibility tests in *Brucella* spp. strains (b). The y-axis represents the relative frequency studies that tested the antimicrobials and the number above the bars the absolute frequency of studies assessing AMR for this drug.

The individual results for each strain were not available in all studies, making it impossible to measure the drug concentrations necessary to inhibit 50 and 90% (MIC50 and MIC90) of the 3,437 field isolates that had their phenotype evaluated in bench tests. However, the MIC50 and MIC90 results obtained by each study considering the tested antimicrobials can be seen in the supplementary file 4 (S4). The number of studies that tested and found resistance to each drug and antimicrobial class are shown in the in Table 1. The antimicrobial classes to which most of the studies observed resistance were respectively ansamycin (21) and folate pathway (8), for MIC tests, and ansamycin (4) and aminoglycoside (4) for disk diffusion. The *Brucella* spp. reference strains mainly tested in the selected studies were *B. melitensis* 16M (35), *B. melitensis* Ether (12), *B. suis* 1330 (15), *B. abortus* 2309 (11), and *B. abortus* 544 (10) (S4 Appendix).

Table 1: Number of studies that tested and observed resistance to several antimicrobials using minimum inhibitory concentration test and disk-diffusion technique in *Brucella* spp. in the selected papers.

| **Antimicrobials** | | **Minimum inibitory concentration** | | | | **Disk difusion** | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Class** | **Drug** | **N of studies that tested (%)** | | **Number of studies that found resistant strains** | | **N of studies that tested (%)** | | **Number of studies that found resistant strains** | |
| | | **Drugs** | **Class** | **Drugs** | **Class** | **Drugs** | **Class** | **Drugs** | **Class** |
| Aminoglycoside | Amikacin | 2 | 44 | 0 | 4 | 0 | 6 | 0 | 4 |
| | Gentamicin | 18 | | 2 | | 3 | | 2 | |
| | Streptomycin | 24 | | 2 | | 3 | | 2 | |
| Ansamycin | Rifampicin | 35 | 35 | 21 | 21 | 4 | 4 | 4 | 4 |
| Beta-lactam | Amoxicilin-clavulanate | 3 | 4 | 0 | 1 | 1 | 4 | 1 | 3 |
| | Ampicillin-sulbactam | 1 | | 1 | | 1 | | 1 | |
| | Ceftazidime | 0 | | 0 | | 1 | | 0 | |
| | Cloxacilim | 0 | | 0 | | 1 | | 1 | |
| Carbapenem | Imipenem | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 |
| Cephem | Ceftriaxone | 8 | 9 | 2 | 2 | 0 | 2 | 0 | 0 |
| | Cefuroxime | 1 | | 0 | | 1 | | 0 | |
| | Cephradine | 0 | | 0 | | 1 | | 0 | |
| Fluoroquinolone | Ciprofloxacin | 25 | 40 | 4 | 5 | 3 | 5 | 1 | 1 |
| | Enrofloxacin | 0 | | 0 | | 1 | | 0 | |
| | Levofloxacin | 8 | | 1 | | 1 | | 0 | |
| | Moxifloxacin | 4 | | 0 | | 0 | | 0 | |
| | Norfloxacin | 2 | | 0 | | 0 | | 0 | |
| | Sparfloxacin | 1 | | 0 | | 0 | | 0 | |
| Folata pathway | Trimethoprim-sulfamethoxazole | 18 | 18 | 8 | 8 | 3 | 3 | 2 | 2 |
| Glycylcycline | Tigecycline | 9 | 9 | 1 | 1 | 1 | 1 | 1 | 1 |
| Lincosamid | Lincomycin | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| Macrolide | Azithromycin | 7 | 11 | 3 | 4 | 0 | 2 | 0 | 2 |
| | Clarithromycin | 1 | | 0 | | 0 | | 0 | |
| | Erythromycin | 3 | | 1 | | 2 | | 2 | |
| Penicicilin | Ampicillin | 3 | 6 | 1 | 1 | 2 | 4 | 2 | 3 |
| | Ofloxacin | 3 | | 0 | | 1 | | 0 | |
| | Penicillim G | 0 | | 0 | | 1 | | 1 | |
| Peptide | Polymixyn B | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 1 |
| | Vancomycin | 0 | | 0 | | 1 | | 1 | |
| Phenicol | Chloramphenicol | 4 | 4 | 0 | 0 | 2 | 2 | 2 | 2 |
| Sulfonamide | Co-trimoxazole | 7 | 7 | 3 | 3 | 0 | 0 | 0 | 0 |
| Tetracycline | Doxycycline | 26 | 43 | 0 | 2 | 1 | 4 | 1 | 3 |
| | Minocycline | 2 | | 0 | | 0 | | 0 | |
| | Tetracycline | 15 | | 2 | | 3 | | 2 | |

**Genotype investigation**

The genotypic evaluation of drug resistance in *Brucella* spp. was performed by 34 studies, totaling 49 different analyzed genes. The most analyzed genes for the occurrence of mutations were *rpoB* (17), *gyrA* (10), *gyrB* (7) and *parC* (5) and the class with most assessed resistance genes were fluoroquinolones (24) and rifampin (17) (Table 1) Detailed information on all the assessed genes are in the Table 1 and additional file 4 (S4 Appendix).

Table 2: Genetic targets for antimicrobial resistance in *Brucella* spp. investigated by the articles selected by this review.

| Gene | N of studies | Product | Mechanism | Antimicrobial class | Author |
|---|---|---|---|---|---|
| *rpoB* | 17 | RNA polymerase β-subunit | Antimicrobial target alteration and antimicrobial target replacement | Ansamycin | (ALJANAZREH; ALZATARI; TAMIMI; ALSAAFEEN *et al.*, 2021; BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021; DADAR; BAZRGARI; GAROSI; HASSAN, 2021; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018; KHAN; MELZER; SAYOUR; SHELL *et al.*, 2021; KHAN; SHELL; MELZER; SAYOUR *et al.*, 2019; LIU; CAO; WANG; PIAO *et al.*, 2019; LIU; DI; WANG; LIU *et al.*, 2018; MARIANELLI; CIUCHINI; TARANTINO; PASQUALI *et al.*, 2004; MARIANELLI; GRAZIANI; SANTANGELO; XIBILIA *et al.*, 2007; MIRJAVADI; KARIMI; AZIMI; GHANAIEE *et al.*, 2020; SANDALAKIS; PSAROULAKI; DE BOCK; CHRISTIDOU *et al.*, 2012; SAYAN; YUMUK; BILENOGLU; ERDENLIG *et al.*, 2009; SAYAN; YUMUK; DUNDAR; BILENOGLU *et al.*, 2008; VALDEZATE; NAVARRO; MEDINA-PASCUAL; CARRASCO *et al.*, 2009; YANG; PIAO; MAO; PANG *et al.*, 2020; YANG; WU; LIU; TIAN *et al.*, 2020) |
| *gyrA* | 10 | DNAgyrase subunit A | Antimicrobial target alteration | Fluoroquinolone | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018; KHAN; MELZER; SAYOUR; SHELL *et al.*, 2021; KHAN; SHELL; MELZER; SAYOUR *et al.*, 2019; LAZARO; RODRIGUEZ-TARAZONA; GARCIA-RODRIGUEZA; MUNOZ-BELLIDO, 2009; MIRJAVADI; KARIMI; AZIMI; GHANAIEE *et al.*, 2020; RAVANEL; GESTIN; MAURIN, 2009; TARAZONA; RODRIGUEZ; BELLIDO, 2014; TURKMANI; PSAROULAKI; CHRISTIDOU; CHOCHLAKIS *et al.*, 2008; VALDEZATE; NAVARRO; MEDINA-PASCUAL; CARRASCO *et al.*, 2009) |
| *gyrB* | 7 | DNA gyrase subunit B | Antimicrobial target alteration | Fluoroquinolone | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018; KHAN; MELZER; SAYOUR; SHELL *et al.*, 2021; KHAN; SHELL; MELZER; SAYOUR *et al.*, 2019; LAZARO; RODRIGUEZ-TARAZONA; GARCIA-RODRIGUEZA; MUNOZ-BELLIDO, 2009; RAVANEL; GESTIN; MAURIN, 2009; VALDEZATE; NAVARRO; MEDINA-PASCUAL; CARRASCO *et al.*, 2009) |
| *parC* | 5 | Topoisomerase IV subunit C | Antimicrobial target alteration | Fluoroquinolone | (JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018; LAZARO; RODRIGUEZ-TARAZONA; GARCIA-RODRIGUEZA; MUNOZ-BELLIDO, 2009; MIRJAVADI; KARIMI; AZIMI; GHANAIEE *et al.*, 2020; RAVANEL; GESTIN; MAURIN, 2009; VALDEZATE; NAVARRO; MEDINA-PASCUAL; CARRASCO *et al.*, 2009) |
| *parE* | 3 | DNA topoisomerase IV subunit B | Antimicrobial target alteration | Fluoroquinolone | (JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018; RAVANEL; GESTIN; MAURIN, 2009; VALDEZATE; NAVARRO; MEDINA-PASCUAL; CARRASCO *et al.*, 2009) |
| *BepC* | 3 | Outer membrane efflux protein | Antimicrobial efflux | Multiple classes | (MARTIN; POSADAS; CARRICA; CRAVERO *et al.*, 2009; POSADAS; MARTIN; GARCIA; SPERA *et al.*, 2007; WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021) |
| *BepE* | 2 | Efflux pump membrane transporter | Antimicrobial efflux | Multiple classes | (MARTIN; POSADAS; CARRICA; CRAVERO *et al.*, 2009; WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021) |
| *BepG* | 2 | Efflux pump membrane transporter | Antimicrobial efflux | Multiple classes | (MARTIN; POSADAS; CARRICA; CRAVERO *et al.*, 2009; WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021) |
| *EF-tuf* | 2 | Elongation factor Tu | Antimicrobial target alteration | Macrolide | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021; HALLING; JENSEN, 2006) |
| *folA* | 2 | Dihydrofolate reductase | Antimicrobial target replacement | Sulfonamides and Diaminopyrimidines | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018) |
| *folP* | 2 | Dihydrofolate reductase | Antimicrobial target replacement | Sulfonamides and Diaminopyrimidines | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018) |
| *mprF* | 2 | Phosphatidylglycerol lysyltransferase | Antimicrobial target alteration | Peptide | (KHAN; MELZER; SAYOUR; SHELL *et al.*, 2021; WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021) |
| *23S* RNA | 2 | 23S ribosomal RNA | Antimicrobial target alteration | Macrolide | (HALLING; JENSEN, 2006; JIANG; MAO; ZHAO; LI *et al.*, 2010) |

| | | | | | |
|---|---|---|---|---|---|
| *aph(3')-IIa* | 2 | Aminoglycoside 3'-O-phosphotransferase | Antimicrobial inactivation | Aminoglycoside | (KHAZAEI; NAJAFI; PIRANFAR; MIRNEJAD, 2016; YANG; WANG; LI; CHEN *et al.*, 2022) |
| *aac*(3'6') | 1 | Aminoglycoside N-acetyltransferase | Antimicrobial inactivation | Aminoglycoside and Fluoroquinolone | (KHAZAEI; NAJAFI; PIRANFAR; MIRNEJAD, 2016) |
| *aadA* | 1 | Aminoglycoside adenylyl-transferase-A1 | Antimicrobial inactivation | Aminoglycoside | (MIRJAVADI; KARIMI; AZIMI; GHANAIEE *et al.*, 2020) |
| *AapJMQ* | 1 | L-amino acid ABC transporter | Antimicrobial efflux | Fluoroquinolone | (YANG; WU; LIU; TIAN *et al.*, 2020) |
| *alr* | 1 | Alanine racemase | Antimicrobial target alteration | Cycloserine | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *ant*(3') | 1 | O-nucleotidyltransferase | Antimicrobial inactivation | Aminoglycoside | (KHAZAEI; NAJAFI; PIRANFAR; MIRNEJAD, 2016) |
| *bepD* | 1 | Efflux pump periplasmic linker | Antimicrobial efflux | Multiple classes | (WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021) |
| *bepF* | 1 | Efflux pump periplasmic linker | Antimicrobial efflux | Multiple classes | (WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021) |
| *bepR* | 1 | HTH-type transcriptional repressor | Regulator modulating expression of Antimicrobial resistance genes | Multiple classes | (MARTIN; POSADAS; CARRICA; CRAVERO *et al.*, 2009) |
| *norM* | 1 | Probable multidrug resistance protein NorM | Antimicrobial efflux | Multiple classes | (BRAIBANT; GUILLOTEAU; ZYGMUNT, 2002) |
| *catB* | 1 | Chloramphenicol acetyltransferase | Antimicrobial inactivation | Phenicol | (KHAZAEI; NAJAFI; PIRANFAR; MIRNEJAD, 2016) |
| *DacC* | 1 | Penicillin-binding protein 6 | Antimicrobial target alteration | Penicillin | (YANG; WU; LIU; TIAN *et al.*, 2020) |
| *ddl* | 1 | D-alanine--D-alanine ligase | Antimicrobial target alteration | Cycloserine | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *dfr* | 1 | Trimethoprim resistant dihydrofolate reductase | Antimicrobial target replacement | Diaminopyrimidine | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *fabl* | 1 | Enoyl-[acyl-carrier-protein] reductase [NADH] | Antimicrobial target replacement | Isoniazid | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *FosA* | 1 | Glutathione transferase | Antimicrobial inactivation | Phosphonic acid | (PARTRIDGE; HALL, 2005) |
| *FosB* | 1 | Metallothiol transferase | Antimicrobial inactivation | Phosphonic acid | (PARTRIDGE; HALL, 2005) |
| *fosX* | 1 | Fosfomycin resistance protein | Antimicrobial inactivation | Phosphonic acid | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *gidB* | 1 | Ribosomal RNA small subunit methyltransferase G | Antimicrobial target alteration | Aminoglycoside | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *gdpD* | 1 | Glycerolphosphodiesterase | Antimicrobial target alteration | Peptide | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *inhA* | 1 | Enoyl-acyl carrier reductase | Antimicrobial target alteration | Isoniazid | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *Iso*-tRNA | 1 | Isoleucyl-tRNA synthetase | Antimicrobial target alteration | Monoxycarbolic acid | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *kasA* | 1 | Ketoacyl acyl carrier protein synthase | Antimicrobial target alteration | Isoniazid | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |

| | | | | | |
|---|---|---|---|---|---|
| *macA* | 1 | Macrolide export protein | Antimicrobial efflux | Macrolide | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *macB* | 1 | Macrolide export ATP-binding/permease protein | Antimicrobial efflux | Macrolide | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *murA* | 1 | UDP-N-acetylglucosamine 1-carboxyvinyltransferase | Antimicrobial target alteration | Phosphonic acid | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *oxyR* | 1 | Transcription factor OxyR | Regulator modulating expression of antimicrobial resistance genes | Isoniazid | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *pgsA* | 1 | Phosphatidylglycerophosphate synthetase | Antimicrobial target alteration | Peptide | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *potA* | 1 | Spermidine/putrescine import ATP-binding protein | Regulator modulating expression of antimicrobial resistance genes | Ansamycin | (YANG; WU; LIU; TIAN *et al.*, 2020) |
| *rho* | 1 | Transcription termination factor Rho | Antimicrobial target alteration | Multiple classes | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *rpID* | 1 | 50S Ribosomal RNA | Antimicrobial target alteration | Macrolide | (HALLING; JENSEN, 2006) |
| *rpoC* | 1 | DNA-directed RNA polymerase β-subunit | Antimicrobial target alteration | Peptide | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *S10p* | 1 | SSU ribosomal protein S10p | Antimicrobial target alteration | Tetracycline | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *S12p* | 1 | SSU ribosomal protein S12p | Antimicrobial target alteration | Aminoglycoside | (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021) |
| *Tet*(WLPSBACOM) | 1 | Tetracycline resistance protein | Antimicrobial efflux | Tetracycline | (KHAZAEI; NAJAFI; PIRANFAR; MIRNEJAD, 2016) |
| *VirB*(1-12) | 1 | Type IV secretion system protein | Antimicrobial intake alteration | Ansamycin | (YANG; WU; LIU; TIAN *et al.*, 2020) |

Among the 49 genes tested for association with AMR in *Brucella* spp., only 15 seemed to be associated with the occurrence of resistance. The relation between AMR in *Brucella* spp. and those genes was established either by the comparison of mutations among resistant and susceptible strains, or through the construction of mutants with subsequent verification of changes in the MIC. The genes reported as probably associated with the occurrence of drug resistance in *Brucella* spp. are shown in Table 3.

Table 3: Genetic targets probably associated with antimicrobial resistance in *Brucella* spp. in the articles selected by this review.

| Gene | Product | Mechanism | Antimicrobial class | Author |
|---|---|---|---|---|
| *aadA* | Aminoglycoside adenylyl-transferase-A1 | Antimicrobial inactivation | Aminoglycoside | (MIRJAVADI et al., 2020) |
| *aph(3')-IIa* | Aminoglycoside 3'-O-phosphotransferase | Antimicrobial inactivation | Aminoglycoside | (YANG; WANG; LI; CHEN *et al.*, 2022) |
| *rpoB* | RNA polymerase β-subunit | Antimicrobial target alteration and antimicrobial target replacement | Ansamycin | (MIRJAVADI et al., 2020, YANG et al., 2020b, KHAN et al., 2021; MARIANELLI et al., 2004; SANDALAKIS et al., 2012) |
| *parC* | Topoisomerase IV subunit C | Antimicrobial target alteration | Fluoroquinolone | (MIRJAVADI et al., 2020) |
| *gyrB* | DNA gyrase subunit B | Antimicrobial target alteration | Fluoroquinolone | (KHAN et al., 2019; KHAN et al., 2021) |
| *gyrA* | DNAgyrase subunit A | Antimicrobial target alteration | Fluoroquinolone | (KHAN et al., 2019; KHAN et al., 2021; RAVANEL et al., 2009; VALDEZATE et al., 2009; LAZARO et al., 2009; TARAZONA et al., 2014; TURKMANI et al., 2008) |
| *23S RNA* | 23S ribosomal RNA | Antimicrobial target alteration | Macrolide | (JIANG et al., 2010) |
| *bepR* | HTH-type transcriptional repressor | Regulator modulating expression of antimicrobial resistance genes | Multiple classes | (MARTIN et al., 2009) |
| *norM* | Probable multidrug resistance protein NorM | Antimicrobial efflux | Multiple classes | (BRAIBANT et al., 2002) |
| *BepC* | Outer membrane efflux protein | Antimicrobial efflux | Multiple classes | (MARTIN et al., 2009; POSADAS et al., 2007) |
| *BepE* | Efflux pump membrane transporter | Antimicrobial efflux | Multiple classes | (MARTIN et al., 2009) |
| *BepG* | Efflux pump membrane transporter | Antimicrobial efflux | Multiple classes | (MARTIN et al., 2009) |
| *DacC* | Penicillin-binding protein 6 | Antimicrobial target alteration | Penicillin | (YANG et al., 2020b) |
| *potA* | Spermidine/putrescine import ATP-binding protein | Regulator modulating expression of Antimicrobial resistance genes | Ansamycin | (YANG et al., 2020b) |
| *VirB(1-12)* | Type IV secretion system protein | Antimicrobial intake alteration | Ansamycin | (YANG et al., 2020b) |

## DISCUSSION

Understanding and monitoring the mechanisms of antimicrobial resistance are priorities at both global and individual (patient care) levels, considering the importance of design strategies that minimize the impact of this problem (BOOLCHANDANI; D'SOUZA; DANTAS, 2019). The emergence of resistance and multidrug resistance in *Brucella* spp. has been described only in the last two decades, and thereby the information available on its frequency and on antimicrobials and genetic determinants involved in this occurrence are still poorly known, which poses a challenge for the implementation of control and prevention measures in this regard. In this context, the goal of this systematic review was to consolidate and discuss the main relevant aspects in the investigation of this phenomenon in *Brucella* spp. available so far. Our results showed that the E-test (28) was technique most used by the studies that investigated the occurrence of AMR in *Brucella* spp., which mainly observed resistance to rifampicin, trimethoprim-sulfamethoxazole and ciprofloxacin, all antimicrobial recommended to treat human brucellosis. Moreover, the results on the genetic determinants of AMR also showed *gyrA* (7), *rpoB* (5) and *bepC* (2), as the genes most associated with the occurrence of drug resistance.

The phenomenon of resistance in *Brucella* spp. has attracted the attention of the scientific community due to its growing occurrence and impact in the treatment of the disease, which requires a combination of drugs for several weeks and has high rate of failure and side effects (YOUSEFI-NOORAIE; MORTAZ-HEJRI; MEHRANI; SADEGHIPOUR, 2012). Lately, a systematic review (WARETH; DADAR; ALI; HAMDY *et al.*, 2022) that aimed to evaluate the therapeutic protocols implemented in the treatment of human brucellosis in Middle Eastern and North Africa pointed out the lack of standardization of phenotypic methods used for measuring drug susceptibility in *Brucella* spp., as observed in the present study. Indeed, since *Brucella* spp. is a fastidious intracellular organism, it would be essential to have specific protocols that should be strictly followed to determine its antimicrobial susceptibility pattern, which could guide the establishment of therapeutic protocols. Nevertheless, the international standards for susceptibility testing of bacteria (CLSI/NCCLS and EUCAST) does not have guidelines specific for *Brucella* spp., and usually recommends the use slow-growing microorganisms manuals, with instructions for inoculum concentration, use of quality control strains, temperature and incubation time, which many researchers still not follow or report. The absence of these critical information on how the tests are performed impair the reliability of the

results observed in some studies. Hence, in this systematic review, we established a minimum quality criteria that should be followed by selected articles, which can be useful for future studies of AMR in *Brucella* spp., contributing to greater uniformity of the methods used and thereby to higher reproducibility of the results in this field. In this scenario, it may be possible to conduct a meta-analysis that will allow the understand in quantitative terms the real occurrence of drug resistance in different *Brucella* species worldwide.

Among the antimicrobial classes investigated in the studies included in this systematic review, those that showed the most resistance in terms of number of publications (not tested strains) were ansamicyn, folate pathway and aminoglycosides. These results are worrisome, since these are the classes that encompass first-choice drugs recommended for the treatment of human brucellosis by WHO guidelines, such as rifampicin, gentamicin, streptomycin and trimethoprim-sulfamethoxazole (CORBEL, 2006). In fact, this has serious implications in the disease treatment, especially considering that most of the analyzed strains in this study were *B. melitensis*, the most virulent species of the genus, isolated from the blood of human patients (causing bacteremia). The occurrence of resistance to these drugs could help to explain the high rates of infection relapses, which can reach up to 40% of cases (FRANCO; MULDER; SMITS, 2007; WARETH; DADAR; ALI; HAMDY *et al.*, 2022).

In addition to phenotypic methods to assess AMR, this review also evaluated the current knowledge on genetic mechanisms of resistance in *Brucella* spp. Since, it is a microorganism with an extremely clonal genome and therefore little prone to lateral gene transfer (SUÁREZ-ESQUIVEL, MARCELA; CHAVES-OLARTE, ESTEBAN; MORENO, EDGARDO; GUZMÁN-VERRI, CATERINA, 2020), it is expected that the occurrence of AMR in this genus does not occur by the presence or absence of genes, but rather at the level of mutations, insertions and deletions (BISWAS; RAOULT; ROLAIN, 2008). Therefore, the simple identification of genes commonly related to AMR in other microorganism in *Brucella* spp., should not be taken as explanation for the mechanisms responsible for this phenotype, as argued by some authors (BOLOTIN; KOVALENKO; MARCHENKO; SOLODIANKIN *et al.*, 2021; KHAZAEI; NAJAFI; PIRANFAR; MIRNEJAD, 2016; PACHECO-MONTEALEGRE; PATIÑO; TORRES; JIMÉNEZ *et al.*, 2017). On the other hand, the identification of single nucleotide polymorphisms (SNPs) through sequencing has been the most used approach in studies so far, but it still has limitations, as it is often not able to explain the reduced susceptibility phenotype of the investigated strains (JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018; LIU; CAO; WANG; PIAO *et al.*, 2019; MIRJAVADI; KARIMI; AZIMI;

GHANAIEE *et al.*, 2020). This occurs because the knowledge about the genes related to AMR in *Brucella* spp. is embryonic, making investigations of only specific fragments of the genome, totally dependent on a previous literature, which is still scarce for *Brucella* spp. (PEREIRA; KATO; ARAÚJO; DA SILVA *et al.*, 2023). Along with SNP detection following gene sequencing, the genome wise association studies (GWAS), which analyzes not only the "target" genes already known in other species, but also their regulatory and promoting regions was the approach that most obtained results in the associations between AMR phenotype and genotype in *Brucella* spp. (YANG; PIAO; MAO; PANG *et al.*, 2020), considering genes hitherto little explored regarding the genetic mechanisms of resistance. In fact, this method has already been applied to detect AMR in other pathogens and seems to be the future in terms of understanding which genes and mutations may be responsible for decreasing the drug susceptibility (BOKMA; VEREECKE; NAUWYNCK; HAESEBROUCK *et al.*, 2021; DIAZ CABALLERO; CLARK; WANG; DONALDSON *et al.*, 2018).

In conclusion, it was observed that the antimicrobials to which *Brucella* spp. showed the highest rates of resistance are the ones recommended to treat human brucellosis (rifampicin, trimethoprim-sulfamethoxazole and ciprofloxacin), which are mainly associated with mutations in the genes *gyrA* and *rpoB*. As recommendations and perspectives, this systematic review stresses the importance of standardizing methods for measuring drug susceptibility in *Brucella* spp. and the potential that investigations of the whole genome has in the discovery of new coding regions and SNPs related to AMR, standing out as an alternative to traditional searches based on the same genes already known in other pathogens.

## SUPPLEMENTARY FILES

https://drive.google.com/drive/folders/1vOtahFFRv7pk0kWVbLRhQkvFIDeOEzKy?usp=drive_link

S1 - PRISMA NMA Checklist of items to include when reporting a systematic review involving a network meta-analysis.

S2 - Extensive overview of search terms.

S3 - Inclusion and exclusion criteria for selection of articles.

S4 - Studies describing AMR in *Brucella* spp. included in this systematic review.

S5 - Studies excluded from this systematic review and reasons.

**ACKNOWLEDGMENTS**

This study was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq [423557/2016-5], Fundação de Amparo à Pesquisa do Estado de Minas Gerais – Fapemig [APQ-00283/2016 and RED-00132-22], Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes and Network of Omics Sciences (RECOM). DO'C received a travel grant from CAPES/PrInt/UFLA [52/2019]. CRP received a scholarship from CAPES/PrInt/UFLA [01/2020]. EMSD is also grateful to CNPq for her fellowship.

## REFERENCES

ABDEL-MAKSOUD, M.; HOUSE, B.; WASFY, M.; ABDEL-RAHMAN, B. *et al.* In vitro antibiotic susceptibility testing of *Brucella* isolates from Egypt between 1999 and 2007 and evidence of probable rifampin resistance. **Ann Clin Microbiol Antimicrob**, 11, n. 24, p. 11-24, 2012.

ALISKAN, H.; CAN, F.; DEMIRBILEK, M.; COLAKOGLU, S. *et al.* Determining in vitro synergistic activities of tigecycline with several other antibiotics against *Brucella melitensis* using checkerboard and time-kill assays. **J Chemother**, 21, n. 1, p. 24-30, 2009.

ALJANAZREH, B.; ALZATARI, K.; TAMIMI, A.; ALSAAFEEN, M. H. *et al.* Brucellosis re-emergence after a decade of quiescence in Palestine, 2015–2017: A seroprevalence and molecular characterization study. **Transbound Emerg Dis**, 2021 2021.

ASADI, F. T.; HASHEMI, S. H.; YOUSEF ALIKHANI, M.; MOGHIMBEIGI, A. *et al.* Clinical and diagnostic aspects of brucellosis and antimicrobial susceptibility of *Brucella* isolates in Hamedan, Iran. **Jpn J Infect Dis**, 70, n. 3, p. 235-238, 2017 2017.

BAYKAM, N.; ESENER, H.; ERGÖNÜL, O.; EREN, S. *et al.* In vitro antimicrobial susceptibility of *Brucella* species. **Int J Antimicrob Agents**, 23, n. 4, p. 405-407, 2004 2004.

BISWAS, S.; RAOULT, D.; ROLAIN, J. M. A bioinformatic approach to understanding antibiotic resistance in intracellular bacteria through whole genome analysis. **Int J Antimicrob Agents**, 32, n. 3, p. 207-220, Sep 2008.

BOKMA, J.; VEREECKE, N.; NAUWYNCK, H.; HAESEBROUCK, F. *et al.* Genome-wide association study reveals genetic markers for antimicrobial resistance in Mycoplasma bovis. **J Microbiology Spectrum**, 9, n. 2, p. e00262-00221, 2021.

BOLOTIN, V.; KOVALENKO, G.; MARCHENKO, N.; SOLODIANKIN, O. *et al.* Complete Genome Sequence of *Brucella abortus* 68, Isolated from Aborted Fetal Sheep in Ukraine. **Microbiol Resour Announc**, 10, n. 10, 2021 2021.

BOOLCHANDANI, M.; D'SOUZA, A. W.; DANTAS, G. Sequencing-based methods and resources to study antimicrobial resistance. **Nature Reviews Genetics**, p. 1, 2019.

BRAIBANT, M.; GUILLOTEAU, L.; ZYGMUNT, M. S. Functional characterization of *Brucella melitensis NorMI*, an efflux pump belonging to the multidrug and toxic compound extrusion family. **Antimicrob Agents Chemother**, 46, n. 9, p. 3050-3053, 2002 2002.

CORBEL, M. J. **Brucellosis in humans and animals**. World Health Organization, 2006. 9241547138.

DADAR, M.; ALAMIAN, S.; BRANGSCH, H.; ELBADAWY, M. *et al.* Determination of virulence-associated genes and antimicrobial resistance profiles in *Brucella* isolates recovered from humans and animals in Iran Using NGS technology. **J Pathogens**, 12, n. 1, p. 82, 2023.

DADAR, M.; BAZRGARI, N.; GAROSI, G. A.; HASSAN, S. Investigation of mutations in the rifampin-resistance-determining region of the *rpoB* gene of *Brucella melitensis* by gene analysis. **Jundishapur Journal of Microbiology** 14, n. 2, 2021 2021.

DEAN, A. S.; CRUMP, L.; GRETER, H.; HATTENDORF, J. *et al.* Clinical manifestations of human brucellosis: a systematic review and meta-analysis. **PLoS Negl Trop Dis**, 6, n. 12, p. e1929, 2012.

DESHMUKH, A.; HAGEN, F.; AL SHARABASI, O.; ABRAHAM, M. *et al.* In vitro antimicrobial susceptibility testing of human *Brucella melitensis* isolates from Qatar between 2014-2015. **BMC Microbiol**, 15, p. 121, 2015.

DIAZ CABALLERO, J.; CLARK, S. T.; WANG, P. W.; DONALDSON, S. L. *et al.* A genome-wide association analysis reveals a potential role for recombination in the evolution of antimicrobial resistance in *Burkholderia multivorans*. **PLoS pathogens**, 14, n. 12, p. e1007453, 2018.

ELBEHIRY, A.; ALDUBAIB, M.; AL RUGAIE, O.; MARZOUK, E. *et al.* Proteomics-based screening and antibiotic resistance assessment of clinical and sub-clinical *Brucella* species: An evolution of brucellosis infection control. **PLoS One**, 17, n. 1, p. e0262551, 2022.

FRANCO, M. P.; MULDER, M.; SMITS, H. L. Persistence and relapse in brucellosis and need for improved treatment. **Transactions of The Royal Society of Tropical Medicine and Hygiene**, 101, n. 9, p. 854-855, 2007.

HALLING, S. M.; JENSEN, A. E. Intrinsic and selected resistance to antibiotics binding the ribosome: analyses of *Brucella 23S rrn*, L4, L22, EF-Tu1, EF-Tu2, efflux and phylogenetic implications. **BMC Microbiol.**, 6, p. 16, 2006.

HASHIM, R.; AHMAD, N.; MOHAMED ZAHIDI, J.; TAY, B. Y. *et al.* Identification and in vitro antimicrobial susceptibility of *Brucella* species isolated from human brucellosis. **Int J Microbiol**, 2014, p. 5, 2014 2014.

HUPE, M. EndNote X9. **Journal of Electronic Resources in Medical Libraries**, 16, n. 3-4, p. 117-119, 2019.

JIANG, H.; MAO, L. L.; ZHAO, H. Y.; LI, L. Y. *et al.* MLVA typing and antibiotic susceptibility of *Brucella* human isolates from Liaoning, China. **Trans R Soc Trop Med Hyg**, 104, n. 12, p. 796-800, Dec 2010.

JOHANSEN, T. B.; SCHEFFER, L.; JENSEN, V. K.; BOHLIN, J. *et al.* Whole-genome sequencing and antimicrobial resistance in *Brucella melitensis* from a Norwegian perspective. **Sci Rep**, 8, n. 1, p. 1-9, 2018.

KHAN, A. U.; MELZER, F.; SAYOUR, A. E.; SHELL, W. S. *et al.* Whole-Genome Sequencing for Tracing the Genetic Diversity of *Brucella abortus* and *Brucella melitensis* Isolated from Livestock in Egypt. **Pathogens**, 10, n. 6, p. 20, 2021.

KHAN, A. U.; SHELL, W. S.; MELZER, F.; SAYOUR, A. E. *et al.* Identification, genotyping and antimicrobial susceptibility testing of *Brucella* spp. isolated from livestock in Egypt. **Microorganisms**, 7, n. 12, p. 603, 2019.

KHAZAEI, Z.; NAJAFI, A.; PIRANFAR, V.; MIRNEJAD, R. Microarray-based long oligonucleotides probe designed for *Brucella* spp. detection and identification of antibiotic susceptibility pattern. **Electron Physician**, 8, n. 4, p. 2297-2303, 2016.

KIIZA, D.; DENAGAMAGE, T.; SERRA, R.; MAUNSELL, F. *et al.* A systematic review of economic assessments for brucellosis control interventions in livestock populations. **Prev Vet Med**, 213, p. 105878, Apr 2023.

LAINE, C. G.; SCOTT, H. M.; ARENAS-GAMBOA, A. M. Human brucellosis: Widespread information deficiency hinders an understanding of global disease frequency. **PLOS Neglected Tropical Diseases**, 16, n. 5, p. e0010404, 2022.

LAZARO, F. G.; RODRIGUEZ-TARAZONA, R. E.; GARCIA-RODRIGUEZA, J. A.; MUNOZ-BELLIDO, J. L. Fluoroquinolone-resistant *Brucella melitensis* mutants obtained in vitro. **Int J Antimicrob Agents**, 34, n. 3, p. 252-254, 2009.

LIU, Z. G.; CAO, X. A.; WANG, M.; PIAO, D. R. *et al.* Whole-Genome Sequencing of a *Brucella melitensis* Strain (BMWS93) Isolated from a Bank Clerk and Exhibiting Complete Resistance to Rifampin. **Microbiol Resour Announc**, 8, n. 33, p. 2, 2019.

LIU, Z. G.; DI, D. D.; WANG, M.; LIU, R. H. *et al.* In vitro antimicrobial susceptibility testing of human *Brucella melitensis* isolates from Ulanqab of Inner Mongolia, China. **BMC Infect Dis**, 18, n. 1, 2018.

MAGALHÃES NETO, A.; CORÇÃO, G.; DASSO, M. G.; KEID, L. B. *et al.* Antimicrobial susceptibility of *Brucella* spp. isolated in Brazil. **Revista de Patologia Tropical**, 43, n. 2, p. 163-172, 2014.

MARIANELLI, C.; CIUCHINI, F.; TARANTINO, M.; PASQUALI, P. *et al.* Genetic bases of the rifampin resistance phenotype in *Brucella* spp. **J Clin Microbiol.**, 42, n. 12, p. 5439-5443, 2004.

MARIANELLI, C.; GRAZIANI, C.; SANTANGELO, C.; XIBILIA, M. T. *et al.* Molecular epidemiological and antibiotic susceptibility characterization of Brucella isolates from humans in Sicily, Italy. **J Clin Microbiol**, 45, n. 9, p. 2923-2928, 2007.

MARTIN, F. A.; POSADAS, D. M.; CARRICA, M. C.; CRAVERO, S. L. *et al.* Interplay between two RND systems mediating antimicrobial resistance in *Brucella suis*. **J Bacteriol**, 191, n. 8, p. 2530-2540, 2009.

MIRJAVADI, A.; KARIMI, A.; AZIMI, L.; GHANAIEE, R. M. *et al.* Livestock brucellosis: Frequency of mutations causing resistance to fluoroquinolone, streptomycin and rifampin. **Archives Of Pediatric Infectious Diseases**, 8, n. 1, 2020.

OUZZANI, M.; HAMMADY, H.; FEDOROWICZ, Z.; ELMAGARMID, A. Rayyan—a web and mobile app for systematic reviews. **Systematic Reviews**, 5, n. 1, p. 210, 2016.

PACHECO-MONTEALEGRE, M.; PATIÑO, R. E.; TORRES, L.; JIMÉNEZ, S. *et al.* The draft genome of *Brucella abortus* strain Ba col-B012, isolated from a dairy farm in Nariño, Colombia, bring new insights into the epidemiology of biovar 4 strains. **Stand Genomic Sci**, 12, n. 1, 2017 2017.

PAGE, M. J.; MCKENZIE, J. E.; BOSSUYT, P. M.; BOUTRON, I. *et al.* The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. **International Journal of Surgery**, 88, p. 105906, 2021.

PAPPAS, G.; AKRITIDIS, N.; BOSILKOVSKI, M.; TSIANOS, E. Brucellosis. **N Engl J Med**, 352, n. 22, p. 2325-2336, Jun 2 2005.

PARTRIDGE, S. R.; HALL, R. M. Gene cassettes potentially encoding fosfomycin resistance determinants. **Antimicrob Agents Chemother**, 49, n. 2, p. 860-861, 2005.

PAULETTI, R. B.; STYNEN, A. P. R.; MOL, J. P. S.; DORNELES, E. M. S. *et al.* Reduced Susceptibility to Rifampicin and Resistance to Multiple Antimicrobial Agents among *Brucella abortus* Isolates from Cattle in Brazil. **PLoS One**, 10, n. 7, p. e0132532, 2015.

PEREIRA, C. R.; KATO, R. B.; ARAÚJO, F. A.; DA SILVA, A. L. *et al.* Genomic investigation of antimicrobial resistance in *Brucella abortus* strains isolated from cattle in Brazil. **Gene Reports**, 31, p. 101777, 2023.

POSADAS, D. M.; MARTIN, F. A.; GARCIA, J.; SPERA, J. M. *et al.* The TolC homologue of *Brucella suis* is involved in resistance to antimicrobial compounds and virulence. **Infect Immun.**, 75, n. 1, p. 379-389, 2007.

R. **R: A language and environment for statistical computing**. Versão 4.1.2. Vienna, Austria: R Development Core Team, 2023.

RAVANEL, N.; GESTIN, B.; MAURIN, M. In vitro selection of fluoroquinolone resistance in *Brucella melitensis*. **Int J Antimicrob Agents.**, 34, n. 1, p. 76-81, 2009.

SANDALAKIS, V.; PSAROULAKI, A.; DE BOCK, P. J.; CHRISTIDOU, A. *et al.* Investigation of rifampicin resistance mechanisms in *Brucella abortus* using MS-driven comparative proteomics. **J Proteome Res**, 11, n. 4, p. 2374-2385, 2012.

SAYAN, M.; YUMUK, Z.; BILENOGLU, O.; ERDENLIG, S. *et al.* Genotyping of *Brucella melitensis* by *rpoB* gene analysis and re-evaluation of conventional serotyping method. **Japanese Journal of Infectious Diseases**, 62, n. 2, p. 160-163, 2009.

SAYAN, M.; YUMUK, Z.; DUNDAR, D.; BILENOGLU, O. *et al.* Rifampicin resistance phenotyping of *Brucella melitensis* by *rpoB* gene analysis in clinical isolates. **J Chemother**, 20, n. 4, p. 431-435, 2008.

SHEVTSOV, A.; SYZDYKOV, M.; KUZNETSOV, A.; SHUSTOV, A. *et al.* Antimicrobial susceptibility of *Brucella melitensis* in Kazakhstan. **Antimicrob Resist Infect Control**, 6, p. 130, 2017.

SIEVERT, C. **Interactive web-based data visualization with R, plotly, and shiny**. CRC Press, 2020. 0429824203.

SINGH, B. B.; KHATKAR, M. S.; AULAKH, R. S.; GILL, J. P. S. *et al.* Estimation of the health and economic burden of human brucellosis in India. **Preventive Veterinary Medicine**, 154, p. 148-155, 2018.

SUÁREZ-ESQUIVEL, M.; CHAVES-OLARTE, E.; MORENO, E.; GUZMÁN-VERRI, C. *Brucella* genomics: macro and micro evolution. **Int J Mol Sci**, 21, n. 20, p. 7749, 2020.

TARAZONA, E. R.; RODRIGUEZ, J. A. G.; BELLIDO, J. L. M. Emergence of quinolone-resistant, topoisomerase-mutant *Brucella* after treatment with fluoroquinolones in a macrophage experimental infection model. **Enferm Infecc Microbiol Clin**, 33, n. 4, p. 248-252, 2014.

TURKMANI, A.; PSAROULAKI, A.; CHRISTIDOU, A.; CHOCHLAKIS, D. *et al.* In vitro-selected resistance to fluoroquinolones in two *Brucella* strains associated with mutational changes in *gyrA*. **Int J Antimicrob Agents**, 32, n. 3, p. 227-232, 2008.

VALDEZATE, S.; NAVARRO, A.; MEDINA-PASCUAL, M. J.; CARRASCO, G. *et al.* Molecular screening for rifampicin and fluoroquinolone resistance in a clinical population of *Brucella melitensis*. **J Antimicrob Chemother**, 65, n. 1, p. 51-53, 2009.

WARETH, G.; DADAR, M.; ALI, H.; HAMDY, M. E. R. *et al.* The perspective of antibiotic therapeutic challenges of brucellosis in the Middle East and North African countries: Current situation and therapeutic management. **Transbound Emerg Dis**, 69, n. 5, p. e1253-e1268, 2022.

WARETH, G.; EL-DIASTY, M.; ABDEL-HAMID, N. H.; HOLZER, K. *et al.* Molecular characterization and antimicrobial susceptibility testing of clinical and non-clinical *Brucella melitensis* and *Brucella abortus* isolates from Egypt. **One Health**, 13, p. 100255, 2021.

WICKHAM, H. ggplot2. **Wiley interdisciplinary reviews: computational statistics**, 3, n. 2, p. 180-185, 2011.

YANG, X.; PIAO, D.; MAO, L.; PANG, B. *et al.* Whole-genome sequencing of rough *Brucella melitensis* in China provides insights into its genetic features. **Emerg Microbes Infect.**, 9, n. 1, p. 2147-2156, 2020.

YANG, X.; WANG, Y.; LI, J.; CHEN, J. *et al.* Genetic characteristics of an amikacin-resistant *Brucella abortus* strain first isolated from Marmota himalayana. **Microbial Pathogenesis**, 164, 2022.

YANG, X.; WU, T.; LIU, W.; TIAN, G. *et al.* Cell membrane components of *Brucella melitensis* play important roles in the resistance of low-level rifampicin. **PLoS Negl Trop Dis.**, 14, n. 12, p. 1-14, 2020.

YOUSEFI-NOORAIE, R.; MORTAZ-HEJRI, S.; MEHRANI, M.; SADEGHIPOUR, P. Antibiotics for treating human brucellosis. **Cochrane Database Syst Rev**, 10, n. 10, p. Cd007179, 2012.

## CHAPTER 2 - Genomic investigation of antimicrobial resistance in *Brucella abortus* strains isolated from cattle in Brazil

Carine Rodrigues Pereira[1]; Rodrigo Bentes Kato[2]; Fabrício Almeida Araújo[2]; Alessandra Lima da Silva[3]; Roselane Gonçalves dos Santos[3]; Thiago de Jesus Sousa[3]; Raquel Costa Neia[4]; Saulo Britto da Silva[5]; Charles H. D. Williamson[6]; John Gillece[6]; Andrey Pereira Lage[7]; David O'Callaghan[8]; Derek Pickard[9]; Rommel Thiago Juca Ramos[2]; Vasco Ariston de Carvalho Azevedo[3]; Jeffrey T. Foster[6]; Elaine Maria Seles Dorneles[1]#

[1]Faculdade de Zootecnia e Medicina Veterinária, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil

[2]Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Pará, Brazil

[3]Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

[4]Faculdade de Ciências Básicas, Universidade Federal Fluminense, Nova Friburgo, Rio de Janeiro, Brazil

[5]Escola de Ciências da Saúde, Universidade do Vale do Itajaí, Itajaí, Santa Catarina, Brazil

[6]Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, Arizona, United States of America

[7]Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

[8]Bacterial Virulence and Infectious Disease, University of Montpellier, Nimes, France

[9]Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom

# Address correspondence to author: +55 (35) 3829-1712 – elaine.dorneles@ufla.br – Universidade Federal de Lavras Caixa Postal 3037 – CEP: 37200-900 – Lavras – MG – Brazil

### ABSTRACT

The emergence of resistance to antimicrobials among *Brucella* spp. has highlighted the need to understand the genetic determinants responsible for these phenotypes. We investigated the gene targets associated with antimicrobial resistance in 53 genomes of *B. abortus* strains, isolated from cattle in Brazil, with resistance or intermediate susceptibility to antimicrobials. The

genetic diversity of the genomes was evaluated using single-nucleotide polymorphism (SNP) differences among the isolates. We compared 18 genes (*gyrB*, *bepD*, *bepE*, *norMII*, *norMI*, *parC*, *bepC*, *folP*, *gyrA*, *rpoB*, *folA*, *rsmG*, *marR*, *parE*, *mprF*, *oxyR*, *bepF*, and *bepG*) previously described as related to antimicrobial resistance in *Brucella* spp. to those of the reference strain *B. abortus* 2308, which was susceptible to all tested antimicrobials. Eight genes had nonsynonymous mutations, deletions, or stop codons in at least one of the analyzed genomes. However, we did not observe any association between the genetic polymorphisms in the evaluated genes and the antimicrobial resistance phenotypes observed in the Brazilian analyzed strains (46 intermediate susceptible to rifampicin, 6 resistance to at least one antimicrobial and 1 multidrug resistance).

**KEYWORDS:** brucellosis; epidemiology; single-nucleotide polymorphism; multidrug resistance

## INTRODUCTION

Brucellosis is one of the most common anthropozoonosis in the world, that is passed to man from domestic livestock and from wild animals. In animals, it causes placentitis, abortion, epididymitis, and reductions of milk production, while in humans it is characterized by non-specific symptoms, such as malaise, night fever, chills, weight loss and arthralgia (Corbel et al., 2006). Animals are usually not treated for brucellosis, in contrast human infections are treated by at least 6 weeks of a combined antibiotic therapy, typically associating doxycycline with rifampicin or an aminoglycoside (YOUSEFI-NOORAIE; MORTAZ-HEJRI; MEHRANI; SADEGHIPOUR, 2012).

Worldwide public health organizations consider antimicrobial resistance (AMR) among the most concerning threats to global health and food security (WHO, 2016; WOAH, 2016). AMR is not generally considered to be a concern and antimicrobial susceptibility testing is not usually recommended for routine laboratories since it increases the risk of laboratory-acquired infections and requires biosafety level 3 facilities (TRAXLER; LEHMAN; BOSSERMAN; GUERRA *et al.*, 2013), which makes it difficult to monitor resistance. In recent decades, however, strains of the genus *Brucella* that are resistant and multidrug-resistant to antimicrobials commonly used in treating human brucellosis have unexpectedly emerged (ABDEL-MAKSOUD; HOUSE; WASFY; ABDEL-RAHMAN *et al.*, 2012; ASADI; HASHEMI; YOUSEF ALIKHANI; MOGHIMBEIGI *et al.*, 2017; DESHMUKH; HAGEN; AL

SHARABASI; ABRAHAM *et al.*, 2015; HASHIM; AHMAD; MOHAMED ZAHIDI; TAY *et al.*, 2014; JIANG; MAO; ZHAO; LI *et al.*, 2010; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018b; KHAN; SHELL; MELZER; SAYOUR *et al.*, 2019; LIU; DI; WANG; LIU *et al.*, 2018; MAGALHÃES NETO; CORÇÃO; DASSO; KEID *et al.*, 2014; PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015; SHEVTSOV; SYZDYKOV; KUZNETSOV; SHUSTOV *et al.*, 2017; WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021) which may have serious consequences for the management of human brucellosis. Inappropriate treatment regimens are associated with relapse and the development of chronic infections.

Bovine brucellosis occurs worldwide and has an enormous economic impact in many developing regions, including Latin America, Asia, Africa, and the Middle East (FRANC; KRECEK; HÄSLER; ARENAS-GAMBOA, 2018). In Brazil, the infection caused by *Brucella abortus* is endemic in all states, with prevalence ranging from 0.9% [confidence interval (CI) 95% (0.30–2.11)] in Santa Catarina to 30.6% [CI 95% (27.4–34.0)] in Mato Grosso do Sul (FERREIRA NETO; SILVEIRA; ROSA; GONÇALVES *et al.*, 2016). The recent isolation from cattle of *B. abortus* strains with intermediate susceptibility or resistance to antimicrobials of choice for treating human brucellosis in six Brazilian states generated substantial concern since human brucellosis is generally contracted from infected animals (FOUSKIS; SANDALAKIS; CHRISTIDOU; TSATSARIS *et al.*, 2018; KRACALIK; ABDULLAYEV; ASADOV; ISMAYILOVA *et al.*, 2016; LINDAHL; VRENTAS; DEKA; HAZARIKA *et al.*, 2020). Therefore, it is essential to get a better understanding of the drug resistance mechanisms of these strains (PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015).

Identifying and tracking the determinants involved in the emergence of AMR is essential to reduce its impact on human and animal infections and to design strategies to restrain its spread (KHAN; MELZER; SAYOUR; SHELL *et al.*, 2021). However, as *Brucella* spp. is an intracellular bacterium with no evidence of plasmids or other means of horizontal gene transfer (SUÁREZ-ESQUIVEL; CHAVES-OLARTE; MORENO; GUZMÁN-VERRI, 2020), AMR occurs exclusively due to spontaneous mutations in the genome (BISWAS; RAOULT; ROLAIN, 2008). This characteristic makes identifying genetic determinants of resistance challenging, as the presence/absence of genes commonly associated with AMR is not helpful. As a complementary approach to traditional culture-based methods for clinical and surveillance applications, advances in third- and fourth-generation sequencing technologies have enabled the analysis of a large volume of genomic data to obtain a broad view and a better understanding of the genetic determinants related to AMR (KÖSER; ELLINGTON; PEACOCK, 2014).

Therefore, this study aimed to investigate the genetic markers of AMR in 53 *B. abortus* genomes from strains with natural resistance, multidrug resistance, or intermediate sensitivity to rifampicin, isolated from cattle in Brazil between 1977 and 2008, by analyzing genes previously reported as associated with AMR in *Brucella* spp.

## MATERIAL AND METHODS

### Bacterial strains

We analyzed a collection of fifty-three *B. abortus* strains isolated and identified by microbiological and molecular tests from cattle in six Brazilian states from 1977 to 2008 (MINHARRO; MOL; DORNELES; PAULETTI *et al.*, 2013). We tested these 53 strains and *B. abortus* strain 2308 for antimicrobial susceptibility against drugs commonly prescribed for the treatment of human brucellosis: ciprofloxacin, doxycycline, streptomycin, gentamicin, rifampicin, and trimethoprim plus sulfamethoxazole (1 part trimethoprim to 19 parts of sulfamethoxazole) using minimum inhibitory concentration (MIC) tests (PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015). The epidemiological and susceptibility information of the strains is shown in Supplementary Table S1.

### Genome sequencing, assembly, and annotation

The Illumina HiSeq 2500 platform (Illumina, USA) amplifying 2 x 151 base pairs paired-end reads was used for the whole genome sequencing (WGS) of the 53 *B. abortus* strains. The quality of the sequencing products was assessed by the FASTQC program (ANDREWS; KRUEGER; SECONDS-PICHON; BIGGINS *et al.*, 2010). A *de novo* sequence assembly was performed using SPAdes v. 3.13.0 (BANKEVICH; NURK; ANTIPOV; GUREVICH *et al.*, 2012) or Edena v. 3.131028 (BANKEVICH; NURK; ANTIPOV; GUREVICH *et al.*, 2012). The assembly products were ordered in CONTIGuator (GALARDINI; BIONDI; BAZZICALUPO; MENGONI, 2011) using the *B. abortus* strains 9-941, A19, or BAB8416 as references in order to separate the two bacterial chromosomes and enable gap closure in the next step (HALLING; PETERSON-BURCH; BRICKER; ZUERNER *et al.*, 2005; LI; KANG; LIN; JIA *et al.*, 2019; WANG; WANG; SUN; BATEER *et al.*, 2020) (Supplementary Table S1). The gaps were closed automatically using the programs GFinisher (GUIZELINI; RAITTZ; CRUZ; SOUZA *et al.*, 2016) and GAPblaster (DE SÁ; MIRANDA; VERAS; DE MELO *et al.*, 2016) and manually using BLAST (https://blast.ncbi.nlm.nih.gov/).

The completed genomes were subjected to curatorial analysis and coverage in the CLC program (QIAGEN, 2020). PROKKA was used for annotation (SEEMANN, 2014).

**Target gene investigation and phylogenomic tree construction**

Identifying the genes reported as possibly related to antimicrobial resistance was performed based on a systematic search in the literature in seven different databases (CABI, Cochrane, ProQuest, PubMed, Scielo, Scopus, and Web of Science) on April 8, 2022. The search terms used were "(Brucell*) AND (antimicrobial* OR antibiotic* OR ciprofloxacin OR doxycycline OR gentamicin OR streptomycin OR ofloxacin OR rifampicin OR sulfamethoxazole OR trimethoprim) AND (susceptib* OR sensitiv* OR resistan*). The search was registered in PROSPERO (300883) and resulted in 3444 articles (unpublished data), of which 23 presented descriptions of genes investigated as potential causes of resistance in *Brucella* spp.

Each article was individually evaluated, and the genes potentially associated with AMR were identified through the name and had their locus tags tracked in the ".gbff" extension file of the *B. abortus* 2308 genome (GCF_000054005.1), retrieved from the NCBI platform (https://www.ncbi.nlm.nih.gov/). The locus tags of the genes allowed access to the product that each one encodes; therefore, a search was performed on the Comprehensive Antibiotic Resistance Database – CARD for all genes and their respective products to verify which classes of antimicrobials had resistance potential related to mutations or alterations in these nucleotide sequences. Only genes potentially associated with reduced susceptibility to antimicrobial classes for which intermediate susceptibility or resistance phenotypes were observed among the tested strains were selected: quinolones (ciprofloxacin), aminoglycosides (gentamicin and streptomycin), folate inhibitors (sulfa + trimethoprim) and rifampin (rifampicin).

The sequences of the selected genes from the *B. abortus* 2308 genome were compared with those of the 53 *B. abortus* genomes using the NCBI-BLAST program (JOHNSON; ZARETSKAYA; RAYTSELIS; MEREZHUK *et al.*, 2008). Then, the sequences of each genome were grouped by genes and aligned using the Muscle algorithm of the MEGA program version 11x (KUMAR; STECHER; LI; KNYAZ *et al.*, 2018). The mutations, deletions and insertions were obtained by aligning assemblies for each genome to reference genes to confirm SNPs presence/absence and then compared with the phenotype previously established for each strain (PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015). These results were paired with a core-genome SNP phylogeny. SNPs were called by the Northern Arizona SNP Pipeline (NASP) program (SAHL; LEMMER; TRAVIS; SCHUPP *et al.*, 2016) using *B. abortus* 2308

(GCF_000054005.1) as a reference genome; a phylogeny was generated with iqtree, and the phylogeny was then edited using iTOL (https://itol.embl.de/upload.cgi).

**Data availability**

The genomes were deposited in NCBI and compiled into a single Bioproject: PRJNA750793. The Biosamples of each genome can be found in Supplementary Table S1.

## RESULTS

### *Brucella* genomes

The genomes were deposited in the format of two chromosomes without gaps, with an average size of 3,283,155 base pairs, 57.25% guanine-cytosine (GC) content, and 264X coverage. The average number of coding sequences (CDS) resulting from the annotation was 3,124. The mean number of SNPs in the Brazilian genomes compared with the *B. abortus* 2308 genome was 371 ± 392, ranging from 95 to 1239.

### Target genes potentially associated with antimicrobial resistance

We first investigated eighteen genes potentially associated with antimicrobial resistance. They are presented in Table 1, along with their locus tag, encoded product, antimicrobial class potentially involved in resistance, reference of the study from which they were extracted, the genomes that showed some mutation, and the nucleotide and amino acid positions of the polymorphisms. When compared with the *B. abortus* 2308 genome (Table 1), ten genes did not show any mutation, insertion, deletion, or alteration in their copy number: *gyrB* (BAB_RS16510), *bepD* (BAB_RS17470), *bepE* (BAB_RS17475), *norMI* (BAB_RS17685), *folP* (BAB_RS20950), *gyrA* (BAB_RS21295), *folA* (BAB_RS22685), *parE* (BAB_RS29415), *mprF* (BAB_RS29450) and *oxyR* (BAB_RS30365). Six genes showed at least one synonymous mutation identified in some of the genomes: *norMII* (BAB_RS17560), *bepC* (BAB_RS20535), *rpoB* (BAB_RS21960), *rsmG* (BAB_RS25740), *marR* (BAB_RS27565) and *bepF* (BAB_RS30440). Moreover, eight genes exhibited nonsynonymous mutations or stop codons in at least one of the analyzed genomes, whose details related to the mutation position and amino acid change in each gene/protein product can be found in Figure and Table 1. A deletion was observed in *bepG* (BAB_RS30450), and *rpoB* (BAB_RS21960) is duplicated in some genomes of this study. All these events were shown in the whole genome SNP-based phylogenomic tree (Figure 1) and in Table 1.

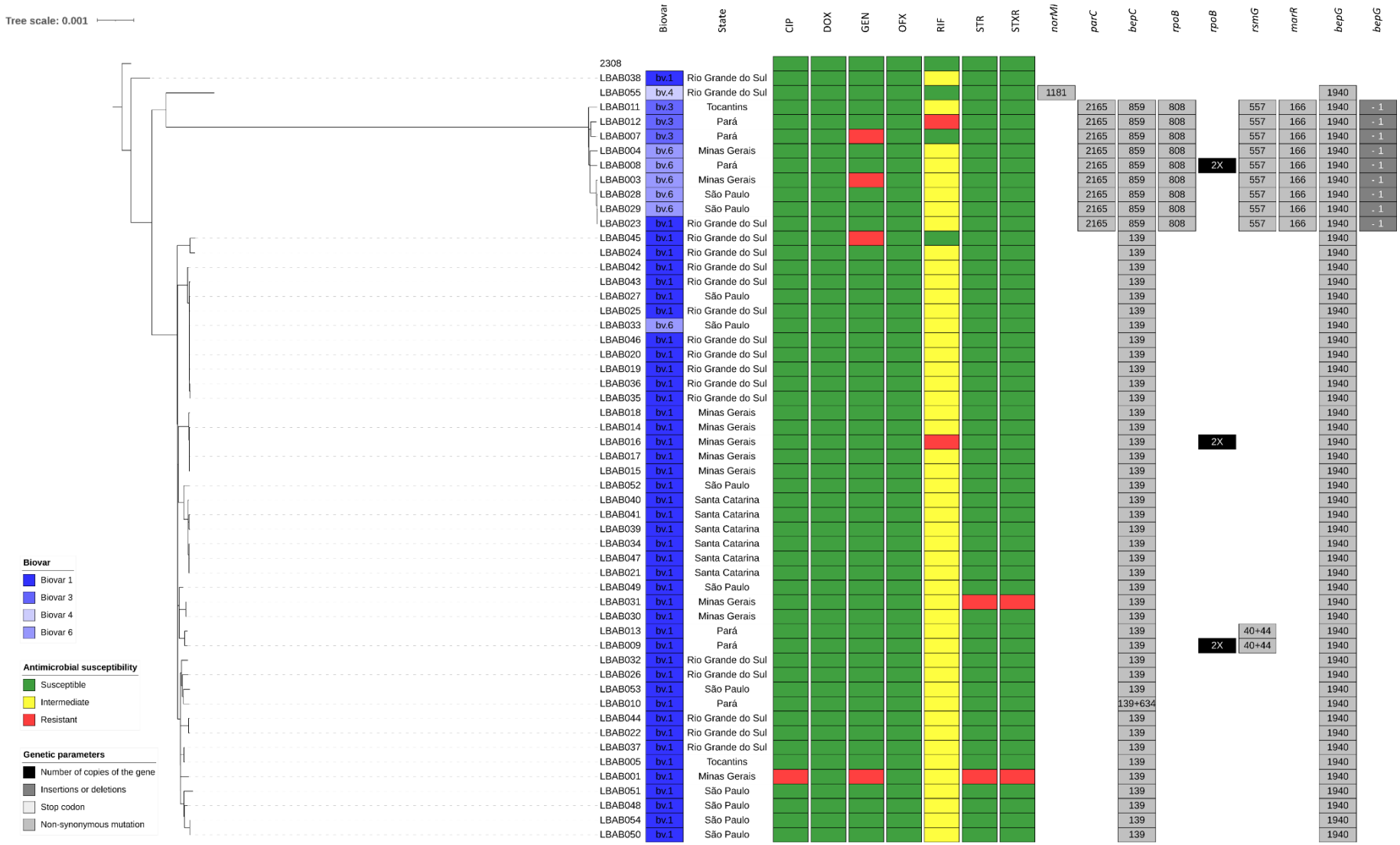


Figure 1: Whole genome SNP-based phylogenetic tree of *B. abortus* strains isolated from cattle in Brazil, with epidemiological State of isolation), microbiological (biovar and antimicrobial susceptibility), and genetic comparison data – SNPs (the number inside the boxes indicates the nucleotide position where the mutation occurred), deletions (- 1 indicates that a nucleotide was deleted) and the number of copies (2x indicates duplicate) of the genes identified by the locus tag of reference strain 2308 – included as outgroup. CIP: ciprofloxacin; DOX: doxycycline; GEN, gentamicin; OFX, ofloxacin; RIF: rifampicin; STR: streptomycin; SXT, sulfamethoxazole + trimethoprim.

**Genomic epidemiology**

A phylogenetic tree was built using the SNPs from the WGS to visualize the relationship of the Brazilian strains isolated from six states that had been classified into different biovars (Figure 1). The tree presented a high clonality among the isolates from Brazil, mainly in the biovar 1 strains. The clade, composed mostly of strains classified as biovars 3 and 6, presented an average of 1,238 SNPs compared with *B. abortus* 2308 and had the longest branch length among the other Brazilian genomes represented in the tree. The biovar 1 and biovar 4 strains (except for LBAB023) showed an average of 194 SNPs compared with the reference genome *B. abortus* 2308. Considering the State from where strains were isolated, Santa Catarina was the only one in which all genomes were grouped.

**DISCUSSION**

In this study, we performed a genomic investigation of *B. abortus* strains isolated from cattle with different levels of susceptibility to the main antibiotics of choice for the treatment of human brucellosis (ciprofloxacin, streptomycin, gentamicin, rifampicin, and trimethoprim plus sulfamethoxazole) by investigating SNPs in genes reported as being responsible for AMR in *Brucella* spp. In addition, a whole-genome SNP analysis was also carried out to better understand Brazil's bovine brucellosis epidemiology in Brazil. In contemporary society, which recognizes the spread of pathogens as a consequence of the globalization process of the modern world, this approach has proven to be a useful strategy to clarify the dynamics of pathogen transmission, upgrade surveillance systems, and minimize the occurrence of new infection cases in humans and animals (DENG; DEN BAKKER; HENDRIKSEN, 2016).

The multidrug resistant (MDR) strain (LBAB001) – resistant to a quinolone (ciprofloxacin), aminoglycosides (gentamicin and streptomycin), and folate inhibitors (sulfa + trimethoprim) – had its complete genome sequenced and analyzed. However, none of the mutations, insertions, or changes in the number of genes previously reported to be associated with resistance to quinolones, aminoglycosides, folate inhibitors, or multidrugs resistance were observed exclusively in this genome. Indeed, although some nonsynonymous mutations were observed in multidrug efflux RND transporter (MARTIN; POSADAS; CARRICA; CRAVERO *et al.*, 2009; WARETH; EL-DIASTY; ABDEL-HAMID; HOLZER *et al.*, 2021) and MATE family efflux transporter genes (BRAIBANT; GUILLOTEAU; ZYGMUNT, 2002; RAVANEL; GESTIN; MAURIN, 2009), no mutations in the analyzed genes showed a possible concordance

compared to the results of different levels of antimicrobial susceptibility. Active efflux of antibiotics interacts with other resistance mechanisms, such as the membrane permeability barrier, enzymatic drug inactivation/modification, and/or alteration/protection of antibiotic targets, significantly increasing resistance levels and profiles (LI; PLÉSIAT; NIKAIDO, 2015). Thus, the isolated analysis of genes may not be able to elucidate this complex mechanism of genetic interactions responsible for the drug resistance phenotype.

In addition to the investigation of genes already well elucidated in the literature as associated with antimicrobial resistance in *Brucella* spp., a recently described gene was also analyzed in this study. The multiple antibiotic resistance regulatory (*marR*) gene was described as responsible for encoding a hypothetical protein in *Brucella* spp., previously characterized in *Escherichia coli* as responsible for encoding the repressor of the mar operon which regulates an efflux system (RAUTHAN; GOEL; KUMAR, 2019). However, the mutations observed in this sequence were more related to a specific clade in the phylogenetic tree than to drug resistance, which is part of a puzzle to better understand gene regulation factors in reducing susceptibility to antibiotics (BOOLCHANDANI; D'SOUZA; DANTAS, 2019). The scientific evidence related to phenotypic tests that may determine the overexpression or repression of genes in *Brucella* spp. through experiments with the construction of mutant strains is still limited compared to other organisms such as *E. coli* regarding resistance to antimicrobials. Thus, a better understanding of the regulation and gene expression in *Brucella* spp. would be important to elucidate the genetic mechanisms of resistance in this genus, which still has 581 protein sequences that remain annotated as "hypothetical" and whose function is unknown (RAUTHAN; GOEL; KUMAR, 2019).

Indeed, the incomplete knowledge of how genetic variants can influence the strain susceptibility to antimicrobials has fomented studies involving two alternative approaches to the conventional search of genes known to be associated with drug resistance: the genomic wide association studies (GWAS) (Lo et al., 2018) and the gene regulatory network analysis (Miryala and Ramaiahand, 2019). The limitation of using the GWAS approach in this study is that the strains are all isolated from Brazil and showed a high similarity to each other, meaning that the SNPs identified in the analysis are more related to the epidemiological context than to the resistance phenotype. To overcome this limitation, a collaborative effort is underway with researchers from different parts of the world aiming to analyze a more diverse collection of strains (which show antimicrobial resistance and the complete genome sequenced). Both the characterization of sequences related to gene expression and regulation and a GWAS approach

are beyond the scope of this study but deserve deep investigation due to the impact of brucellosis as a neglected disease and the importance of surveillance on the spread of drug resistance markers (BAKER; THOMSON; WEILL; HOLT, 2018; FRANC; KRECEK; HÄSLER; ARENAS-GAMBOA, 2018).

Among the tested antimicrobials, the one with the highest occurrence of resistance or intermediate susceptibility was rifampicin (PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015), which is alarming since this drug is one of the two antibiotics recommended by the World Health Organization (WHO) as the first choice for the treatment of human brucellosis. The most common reasons for resistance to rifampicin are mutations in the *rpoB* gene, which encodes the beta subunit of DNA-directed RNA polymerase (MARIANELLI; CIUCHINI; TARANTINO; PASQUALI *et al.*, 2004). An Asp-526-Tyr substitution is associated with the rifampicin-resistant phenotype of the RB51 vaccine strain and substitutions of Asp 526 were also identified in laboratory selected rifampicin-resistant *B. melitensis* mutants, as were substitutions of His 154, His 536 and Ser 541 (MARIANELLI; CIUCHINI; TARANTINO; PASQUALI *et al.*, 2004). All the *rpoB* gene mutations found in the genomes from this study were previously described in the literature (MARIANELLI; CIUCHINI; TARANTINO; PASQUALI *et al.*, 2006); and none of them were related to rifampicin resistance, as confirmed in several other studies (DADAR; BAZRGARI; GAROSI; HASSAN, 2021; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018a; KHAN; SHELL; MELZER; SAYOUR *et al.*, 2019; MARIANELLI; GRAZIANI; SANTANGELO; XIBILIA *et al.*, 2007; VALDEZATE; NAVARRO; MEDINA-PASCUAL; CARRASCO *et al.*, 2009). This gene can have other implications beyond the relationship with the phenotype of resistance to antimicrobials. Due to its high diversity, the *rpoB* gene has been described as a genotyping target to identify phylogenetically related strains (JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018a; MARIANELLI; CIUCHINI; TARANTINO; PASQUALI *et al.*, 2006). In our Brazilian genomes, the SNPs found in this gene do not follow a pattern related to antimicrobial susceptibility but were all found in the most heterogeneous clade of the phylogenetic tree, composed mostly of strains classified as biovars 3 and 6, which reinforces its usefulness as an epidemiological marker (BAZRGARI; GAROOSI; DADAR, 2020; JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018a; MARIANELLI; CIUCHINI; TARANTINO; PASQUALI *et al.*, 2006).

In addition to the mutations observed in the *rpoB* gene, the occurrence of duplications of this sequence in three phylogenetically unrelated genomes that exhibited intermediate

susceptibility or resistance to rifampicin – LBAB008, LBAB009, and LBAB016 – was identified. Duplication of *rpoB* is only associated with resistance to rifampicin when at least one of the sequences has mutations in one of the nucleotides that encode the amino acid residues of the β subunit of RNA polymerase (67), which was not observed for any of the analyzed genomes. However, the existence of extra copies of some genes may be related to their overexpression and thus responsible for the observed resistance phenotype (64).

Resistance to fluoroquinolones is mainly associated with mechanisms controlled by the proteins encoded by *gyrA*, *gyrB*, *parC,* and *parE* (VALDEZATE; NAVARRO; MEDINA-PASCUAL; CARRASCO *et al.*, 2009). In this study, SNPs in these four genes were investigated in all Brazilian strains; however, only one nonsynonymous mutation was observed in the *parC* gene, which is responsible for encoding DNA topoisomerase IV subunit A. The polymorphism observed in amino acid 722 of this gene was found in nine Brazilian *B. abortus* genomes, all from strains susceptible to ciprofloxacin (PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015). However, the same SNP was also described in the genome of the vaccine strain S19 (JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018a), which exhibits a high sensitivity to this antimicrobial (PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015). Therefore, this mutation appears to be not associated with the occurrence of resistance to this quinolone. Resistance to this class of antibiotics can be attributed to multifactorial mechanisms, including changes in fluoroquinolone target enzymes, as investigated in DNA gyrases and topoisomerases IV, and a reduction in the interbacterial drug concentration and changes in drug target enzymes (VALDEZATE; NAVARRO; MEDINA-PASCUAL; CARRASCO *et al.*, 2009).

In addition to contributing to the understanding of genetic markers associated with AMR in *Brucella* spp., to the best of our knowledge, this is the first study to perform a WGS investigation of *B. abortus* strains isolated from cattle in Brazil. The WGS of Brazilian strains and their deposition at NCBI increased the number of complete *B. abortus* genomes available by more than 300%, from 23 to 76. This increase in complete genomes is of great importance since although the cost of high-throughput third-generation sequencing technologies and the development of assembly algorithms have evolved over the years, the deposit of WGS at the level of scaffolds and contigs, such as fragmented and unfinished drafts, still represents the majority of genomes available (LIAO; LIN; LIN, 2015). Among several limitations of performing analyses with unfinished and incomplete genomes is that the gapped regions may contain essential information lost in the assembly process (BOETZER; PIROVANO, 2012).

The phylogenomic analysis showed high clonality among the 53 assessed genomes. Despite the sampling of this study being non-probabilistic and therefore not representative of the reality of the entire country, among the states from which strains were collected, the one that presented the most homogeneous genomes was Santa Catarina; this may be explained by the fact that the state has the lowest prevalence of bovine brucellosis in the country (BAUMGARTEN; VELOSO; GRISI-FILHO; FERREIRA *et al.*, 2016). Currently in the eradication phase of the disease, Santa Catarina has a lower number of circulating *B. abortus* strains and consequently they are more similar to each other at the WGS level. The genomes from *B. abortus* isolated from other states, such as those from Minas Gerais and São Paulo, grouped in different clades, which could be attributed to the wide circulation of the pathogen in Brazilian territory and the high level of clonality of the species. Future studies examining the relationship between epidemiological and genomic information can be carried out using these sequencing data, thereby generating a better understanding of the transmission dynamics of this zoonosis in the country and, consequently, the proposition of more strategic measures for its control. For this to be accomplished, it is important to compare not only the genomes of the Brazilian strains with each other but also with the other genomes of the species publicly available at the NCBI with their respective epidemiological information. Thus, hypotheses regarding transmission dynamics can be posed both about the sources of entry of the pathogen in Brazil and its circulation in the national territory. Such an approach is fundamental in the low prevalence scenarios since the more advanced disease control requires more advanced levels of analysis, with genomic epidemiology being one of the most used strategies for this purpose (JOHANSEN; SCHEFFER; JENSEN; BOHLIN *et al.*, 2018b; KAMATH; DREES; FOSTER; QUANCE *et al.*, 2014; MUÑOZ; MICK; SACCHINI; JANOWICZ *et al.*, 2019).

## CONCLUSION

The mutations found in the genes previously reported as potentially associated with AMR in *Brucella* spp. did not support the phenotype and genotype association among the assessed isolates. Furthermore, from the WGS approach, insights into the genomic epidemiology of *B. abortus* strains isolated in Brazil could be considered a first step toward supporting the control and eradication of bovine brucellosis in the country.

## SUPPLEMENTARY FILES

https://drive.google.com/drive/folders/1vOtahFFRv7pk0kWVbLRhQkvFIDeOEzKy?usp=drive_link

S1 - Epidemiological information of the strains, susceptibility to antimicrobials, genome of reference used to order the assembled contigs and accession number of the strains at GenBank.

## ACKNOWLEDGMENTS

We want to acknowledge Laís Gabrielle Alvarenga Assis for her technical support. This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq [423557/2016-5], Fundação de Amparo à Pesquisa do Estado de Minas Gerais – Fapemig [00283/2016 and RED-00132-22], Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes and Network of Omics Sciences (RECOM). DO'C and CRP received a travel grant from CAPES/PrInt/UFLA. VBIC/CNR *Brucella* was supported by INSERM, Université de Montpellier, and Santé Public France. The Wellcome Trust sponsored DJJP.

## REFERENCES

ABDEL-MAKSOUD, M.; HOUSE, B.; WASFY, M.; ABDEL-RAHMAN, B. *et al.* In vitro antibiotic susceptibility testing of *Brucella* isolates from Egypt between 1999 and 2007 and evidence of probable rifampin resistance. **Ann Clin Microbiol Antimicrob**, 11, n. 24, p. 11-24, 2012.

ADONE, R.; CIUCHINI, F.; MARIANELLI, C.; TARANTINO, M. *et al.* Protective properties of rifampin-resistant rough mutants of *Brucella melitensis*. **Infect Immun**, 73, n. 7, p. 4198-4204, 2005.

ALJANAZREH, B.; ALZATARI, K.; TAMIMI, A.; ALSAAFEEN, M. H. *et al.* Brucellosis re-emergence after a decade of quiescence in Palestine, 2015–2017: A seroprevalence and molecular characterization study. **Transbound Emerg Dis**, 69, n. 4, p. e130-e140, 2022.

ANDREWS, S.; KRUEGER, F.; SECONDS-PICHON, A.; BIGGINS, F. *et al.* **A quality control tool for high throughput sequence data.** Babraham Bioinformatics, 2010.

ASADI, F. T.; HASHEMI, S. H.; YOUSEF ALIKHANI, M.; MOGHIMBEIGI, A. *et al.* Clinical and diagnostic aspects of brucellosis and antimicrobial susceptibility of *Brucella* isolates in Hamedan, Iran. **Jpn J Infect Dis**, 70, n. 3, p. 235-238, 2017.

BAKER, S.; THOMSON, N.; WEILL, F. X.; HOLT, K. E. Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. **Science**, 360, n. 6390, p. 733-738, 2018.

BANKEVICH, A.; NURK, S.; ANTIPOV, D.; GUREVICH, A. A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **J Comput Biol**, 19, n. 5, p. 455-477, 2012.

BAUMGARTEN, K. D.; VELOSO, F. P.; GRISI-FILHO, J. H. H.; FERREIRA, F. *et al.* Prevalence and risk factors for bovine brucellosis in the State of Santa Catarina, Brazil. **Semin Cienc Agrar**, 37, n. 5Supl2, p. 3425, 2016.

BAZRGARI, N.; GAROOSI, G. A.; DADAR, M. Genetic diversity and phylogenetic relationship of clinical isolates of *Brucella melitensis* based on gene polymorphism of β subunit of RNA Polymerase (*rpoB*) gene in Iran. **Iran J Med Microbiol**, 14, n. 5, p. 425-440, 2020.

BISWAS, S.; RAOULT, D.; ROLAIN, J. M. A bioinformatic approach to understanding antibiotic resistance in intracellular bacteria through whole genome analysis. **Int J Antimicrob Agents**, 32, n. 3, p. 207-220, 2008.

BOETZER, M.; PIROVANO, W. Toward almost closed genomes with GapFiller. **Genome Biol**, 13, n. 6, p. R56, 2012/06/25 2012.

BOLOTIN, V.; KOVALENKO, G.; MARCHENKO, N.; SOLODIANKIN, O. *et al.* Complete Genome Sequence of *Brucella abortus* 68, Isolated from Aborted Fetal Sheep in Ukraine. **Microbiol Resour Announc**, 10, n. 10, p. e01436-01420, 2021.

BOOLCHANDANI, M.; D'SOUZA, A. W.; DANTAS, G. Sequencing-based methods and resources to study antimicrobial resistance. **Nat Rev Genet**, 20, n. 6, p. 356-370, 2019.

BRAIBANT, M.; GUILLOTEAU, L.; ZYGMUNT, M. S. Functional characterization of *Brucella melitensis NorMI*, an efflux pump belonging to the multidrug and toxic compound extrusion family. **Antimicrob Agents Chemother**, 46, n. 9, p. 3050-3053, 2002.

DADAR, M.; BAZRGARI, N.; GAROSI, G. A.; HASSAN, S. Investigation of mutations in the rifampin-resistance-determining region of the *rpoB* gene of *Brucella melitensis* by gene analysis. **Jundishapur J Microbiol**, 14, n. 2, p. e115526, 2021.

DE SÁ, P. H. C. G.; MIRANDA, F.; VERAS, A.; DE MELO, D. M. *et al.* GapBlaster—a graphical gap filler for prokaryote genomes. **PLoS One**, 11, n. 5, p. e0155327, 2016.

DENG, X.; DEN BAKKER, H. C.; HENDRIKSEN, R. Genomic epidemiology: whole-genome-sequencing-powered surveillance and outbreak investigation of foodborne bacterial pathogens. **Annu Rev Food Sci Technol**, 7, n. 1, p. 353-374, 2016.

DESHMUKH, A.; HAGEN, F.; AL SHARABASI, O.; ABRAHAM, M. *et al.* In vitro antimicrobial susceptibility testing of human *Brucella melitensis* isolates from Qatar between 2014-2015. **BMC Microbiol**, 15, p. 121, 2015.

FERREIRA NETO, J. S.; SILVEIRA, G. B.; ROSA, B. M.; GONÇALVES, V. S. P. *et al.* Analysis of 15 years of the National Program for the Control and Eradication of Animal Brucellosis and Tuberculosis, Brazil. **Semin Cienc Agrar**, 37, n. 5, suppl.2, p. 3385-3402, 2016.

FOUSKIS, I.; SANDALAKIS, V.; CHRISTIDOU, A.; TSATSARIS, A. *et al.* The epidemiology of Brucellosis in Greece, 2007–2012: a 'One Health' approach. **Trans R Soc Trop Med Hyg**, 112, n. 3, p. 124-135, 2018.

FRANC, K. A.; KRECEK, R. C.; HÄSLER, B. N.; ARENAS-GAMBOA, A. M. Brucellosis remains a neglected disease in the developing world: a call for interdisciplinary action. **BMC Public Health**, 18, n. 1, p. 1-9, 2018.

GALARDINI, M.; BIONDI, E. G.; BAZZICALUPO, M.; MENGONI, A. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. **Source Code Biol Med**, 6, n. 1, p. 11, 2011.

GUIZELINI, D.; RAITTZ, R. T.; CRUZ, L. M.; SOUZA, E. l. M. *et al.* GFinisher: a new strategy to refine and finish bacterial genome assemblies. **Sci Rep**, 6, p. 34963, 2016.

HALLING, S. M.; PETERSON-BURCH, B. D.; BRICKER, B. J.; ZUERNER, R. L. *et al.* Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. **J Bacteriol**, 187, n. 8, p. 2715-2726, 2005.

HASHIM, R.; AHMAD, N.; MOHAMED ZAHIDI, J.; TAY, B. Y. *et al.* Identification and in vitro antimicrobial susceptibility of *Brucella* species isolated from human brucellosis. **Int J Microbiol**, 2014, p. 5, 2014 2014.

JIANG, H.; MAO, L. L.; ZHAO, H. Y.; LI, L. Y. *et al.* MLVA typing and antibiotic susceptibility of *Brucella* human isolates from Liaoning, China. **Trans R Soc Trop Med Hyg**, 104, n. 12, p. 796-800, 2010.

JOHANSEN, T. B.; SCHEFFER, L.; JENSEN, V. K.; BOHLIN, J. *et al.* Whole-genome sequencing and antimicrobial resistance in *Brucella melitensis* from a Norwegian perspective. 8, n. 1, 2018 2018a.

JOHANSEN, T. B.; SCHEFFER, L.; JENSEN, V. K.; BOHLIN, J. *et al.* Whole-genome sequencing and antimicrobial resistance in *Brucella melitensis* from a Norwegian perspective. **Sci Rep**, 8, n. 1, p. 1-9, 2018b.

JOHNSON, M.; ZARETSKAYA, I.; RAYTSELIS, Y.; MEREZHUK, Y. *et al.* NCBI BLAST: a better web interface. **Nucleic Acids Res**, 36, n. suppl_2, p. W5-W9, 2008.

KAMATH, P. L.; DREES, K. P.; FOSTER, J. T.; QUANCE, C. *et al.* Genomics of Brucellosis in Wildlife and Livestock of the Greater Yellowstone Ecosystem. **Intermt J Sci**, 20, n. 4, p. 100-100, 2014.

KHAN, A. U.; MELZER, F.; SAYOUR, A. E.; SHELL, W. S. *et al.* Whole-Genome Sequencing for Tracing the Genetic Diversity of *Brucella abortus* and *Brucella melitensis* Isolated from Livestock in Egypt. **Pathogens**, 10, n. 6, p. 20, 2021.

KHAN, A. U.; SHELL, W. S.; MELZER, F.; SAYOUR, A. E. *et al.* Identification, genotyping and antimicrobial susceptibility testing of *Brucella* spp. isolated from livestock in Egypt. **Microorganisms**, 7, n. 12, p. 603, 2019.

KÖSER, C. U.; ELLINGTON, M. J.; PEACOCK, S. J. Whole-genome sequencing to control antimicrobial resistance. **Trends Genet**, 30, n. 9, p. 401-407, 2014.

KRACALIK, I. T.; ABDULLAYEV, R.; ASADOV, K.; ISMAYILOVA, R. *et al.* Human Brucellosis Trends: Re-emergence and Prospects for Control Using a One Health Approach in Azerbaijan (1983–2009). **Zoonoses Public Health**, 63, n. 4, p. 294-302, 2016.

KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C. *et al.* MEGA X: molecular evolutionary genetics analysis across computing platforms. **Mol Biol Evol**, 35, n. 6, p. 1547, 2018.

LAZARO, F. G.; RODRIGUEZ-TARAZONA, R. E.; GARCIA-RODRIGUEZA, J. A.; MUNOZ-BELLIDO, J. L. Fluoroquinolone-resistant *Brucella melitensis* mutants obtained in vitro. **Int J Antimicrob Agents**, 34, n. 3, p. 252-254, 2009.

LI, X.; KANG, Y.; LIN, L.; JIA, E. *et al.* Genomic Characterization provides new insights for detailed phage-resistant mechanism for *Brucella abortus*. **Front Microbiol**, 10, p. 917, 2019.

LI, X. Z.; PLÉSIAT, P.; NIKAIDO, H. The Challenge of Efflux-Mediated Antibiotic Resistance in Gram-Negative Bacteria. **Clin Microbiol Rev**, 28, n. 2, p. 337-418, 2015/04/01 2015.

LIAO, Y. C.; LIN, S. H.; LIN, H. H. Completing bacterial genome assemblies: strategy and performance comparisons. **Sci Rep**, 5, n. 1, p. 8747, 2015/03/04 2015.

LINDAHL, J. F.; VRENTAS, C. E.; DEKA, R. P.; HAZARIKA, R. A. *et al.* Brucellosis in India: results of a collaborative workshop to define One Health priorities. **Trop Anim Health Prod**, 52, n. 1, p. 387-396, Jan 2020.

LIU, Z. G.; DI, D. D.; WANG, M.; LIU, R. H. *et al.* In vitro antimicrobial susceptibility testing of human *Brucella melitensis* isolates from Ulanqab of Inner Mongolia, China. **BMC Infect Dis**, 18, n. 1, 2018.

MAGALHÃES NETO, A.; CORÇÃO, G.; DASSO, M. G.; KEID, L. B. *et al.* Antimicrobial susceptibility of *Brucella* spp. isolated in Brazil. **Revista de Patologia Tropical**, 43, n. 2, p. 163-172, 2014.

MARIANELLI, C.; CIUCHINI, F.; TARANTINO, M.; PASQUALI, P. *et al.* Genetic bases of the rifampin resistance phenotype in *Brucella* spp. **J Clin Microbiol**, 42, n. 12, p. 5439-5443, 2004 2004.

MARIANELLI, C.; CIUCHINI, F.; TARANTINO, M.; PASQUALI, P. *et al.* Molecular characterization of the *rpoB* gene in *Brucella* species: new potential molecular markers for genotyping. **Microbes Infect**, 8, n. 3, p. 860-865, 2006.

MARIANELLI, C.; GRAZIANI, C.; SANTANGELO, C.; XIBILIA, M. T. *et al.* Molecular epidemiological and antibiotic susceptibility characterization of *Brucella* isolates from humans in Sicily, Italy. **J Clin Microbiol**, 45, n. 9, p. 2923-2928, 2007.

MARTIN, F. A.; POSADAS, D. M.; CARRICA, M. C.; CRAVERO, S. L. *et al.* Interplay between two RND systems mediating antimicrobial resistance in *Brucella suis*. **J Bacteriol**, 191, n. 8, p. 2530-2540, 2009.

MINHARRO, S.; MOL, J. P. S.; DORNELES, E. M. S.; PAULETTI, R. B. *et al.* Biotyping and genotyping (MLVA16) of *Brucella abortus* isolated from cattle in Brazil, 1977 to 2008. **PLoS One**, 8, n. 12, p. e81152, 2013.

MIRJAVADI, A.; KARIMI, A.; AZIMI, L.; GHANAIEE, R. M. *et al.* Livestock brucellosis: Frequency of mutations causing resistance to fluoroquinolone, streptomycin and rifampin. **Arch Pediatr Infect Dis**, 8, n. 1, p. e83557, 2020.

MUÑOZ, P. M.; MICK, V.; SACCHINI, L.; JANOWICZ, A. *et al.* Phylogeography and epidemiology of *Brucella suis* biovar 2 in wildlife and domestic swine. **Vet Microbiol**, 233, p. 68-77, 2019.

PAULETTI, R. B.; STYNEN, A. P. R.; MOL, J. P. S.; DORNELES, E. M. S. *et al.* Reduced Susceptibility to Rifampicin and Resistance to Multiple Antimicrobial Agents among *Brucella abortus* Isolates from Cattle in Brazil. **PLoS One**, 10, n. 7, p. e0132532, 2015.

POSADAS, D. M.; MARTIN, F. A.; GARCIA, J.; SPERA, J. M. *et al.* The *TolC* homologue of *Brucella suis* is involved in resistance to antimicrobial compounds and virulence. **Infect Immun**, 75, n. 1, p. 379-389, 2007.

QIAGEN. **CLC Genomics Workbench**. Versão 20.0. 2020.

RAUTHAN, K.; GOEL, D.; KUMAR, S. Annotation of a hypothetical protein (WP_002969292.1) from *Brucella abortus*. **Bioinformation**, 15, n. 5, p. 315-320, 2019.

RAVANEL, N.; GESTIN, B.; MAURIN, M. In vitro selection of fluoroquinolone resistance in *Brucella melitensis*. **Int J Antimicrob Agents**, 34, n. 1, p. 76-81, 2009.

SAHL, J. W.; LEMMER, D.; TRAVIS, J.; SCHUPP, J. M. *et al.* The Northern Arizona SNP Pipeline (NASP): accurate, flexible, and rapid identification of SNPs in WGS datasets. **bioRxiv**, p. 037267, 2016.

SANDALAKIS, V.; PSAROULAKI, A.; DE BOCK, P. J.; CHRISTIDOU, A. *et al.* Investigation of rifampicin resistance mechanisms in *Brucella abortus* using MS-driven comparative proteomics. **J Proteome Res**, 11, n. 4, p. 2374-2385, 2012.

SAYAN, M.; YUMUK, Z.; DUNDAR, D.; BILENOGLU, O. *et al.* Rifampicin resistance phenotyping of *Brucella melitensis* by *rpoB* gene analysis in clinical isolates. **J Chemother**, 20, n. 4, p. 431-435, 2008.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, 30, n. 14, p. 2068-2069, 2014.

SHEVTSOV, A.; SYZDYKOV, M.; KUZNETSOV, A.; SHUSTOV, A. *et al.* Antimicrobial susceptibility of *Brucella melitensis* in Kazakhstan. **Antimicrob Resist Infect Control**, 6, p. 130, 2017.

SUÁREZ-ESQUIVEL, M.; CHAVES-OLARTE, E.; MORENO, E.; GUZMÁN-VERRI, C. *Brucella* genomics: macro and micro evolution. **Int J Mol Sci**, 21, n. 20, p. 7749, 2020.

TARAZONA, E. R.; RODRIGUEZ, J. A. G.; BELLIDO, J. L. M. Emergence of quinolone-resistant, topoisomerase-mutant *Brucella* after treatment with fluoroquinolones in a macrophage experimental infection model. **Enferm Infecc Microbiol Clin**, 33, n. 4, p. 248-252, 2014.

TRAXLER, R. M.; LEHMAN, M. W.; BOSSERMAN, E. A.; GUERRA, M. A. *et al.* A literature review of laboratory-acquired brucellosis. **J Clin Microbiol**, 51, n. 9, p. 3055-3062, 2013.

TURKMANI, A.; PSAROULAKI, A.; CHRISTIDOU, A.; CHOCHLAKIS, D. *et al.* In vitro-selected resistance to fluoroquinolones in two *Brucella* strains associated with mutational changes in *gyrA*. **Int J Antimicrob Agents**, 32, n. 3, p. 227-232, 2008.

VALDEZATE, S.; NAVARRO, A.; MEDINA-PASCUAL, M. J.; CARRASCO, G. *et al.* Molecular screening for rifampicin and fluoroquinolone resistance in a clinical population of *Brucella melitensis*. **J Antimicrob Chemother**, 65, n. 1, p. 51-53, 2009.

WANG, S.; WANG, W.; SUN, K.; BATEER, H. *et al.* Comparative genomic analysis between newly sequenced *Brucella abortus* vaccine strain A19 and another *Brucella abortus* vaccine S19. **Genomics**, 112, n. 2, p. 1444-1453, 2020.

WARETH, G.; EL-DIASTY, M.; ABDEL-HAMID, N. H.; HOLZER, K. *et al.* Molecular characterization and antimicrobial susceptibility testing of clinical and non-clinical *Brucella melitensis* and *Brucella abortus* isolates from Egypt. **One Health**, 13, p. 100255, 2021 2021.

WHO. Global action plan on antimicrobial resistance. **World Health Organization** 2016.

WOAH. The OIE strategy on antimicrobial resistance and the prudent use of antimicrobials. **World Organization for Animal Health**. Paris, France 2016.

YOUSEFI-NOORAIE, R.; MORTAZ-HEJRI, S.; MEHRANI, M.; SADEGHIPOUR, P. Antibiotics for treating human brucellosis. **Cochrane Database Syst Rev**, n. 10, 2012.

ZAKI, N. A.; SALLOUM, T.; OSMAN, M.; RAFEI, R. *et al.* Typing and comparative genome analysis of *Brucella melitensis* isolated from Lebanon. **FEMS Microbiol Lett**, 364, n. 19, 2017 2017.

# CHAPTER 3 - Comparison of *Brucella abortus* population structure based on genotyping methods with different levels of resolution

Carine R. Pereira[1]; Raquel C. Neia[2]; Saulo B. Silva[3]; Charles H. D. Williamson[4]; John D. Gillece[4]; David O'Callaghan[5]; Jeffrey T. Foster[4]; Izabela R. C. Oliveira[6]; Júlio S. S. B. Filho[6]; Andrey P. Lage[7]; Vasco A. C. Azevedo[8]; Elaine M. S. Dorneles[1*]

[1]Faculdade de Zootecnia e Medicina Veterinária, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil

[2]Faculdade de Ciências Básicas, Universidade Federal Fluminense, Nova Friburgo, Rio de Janeiro, Brazil

[3]Escola de Ciências da Saúde, Univali, Itajaí, Santa Catarina, Brazil

[4]Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, Arizona, USA

[5]Bacterial Virulence and Infectious Disease, University of Montpellier, Nimes, France

[6]Departamento de Estatística, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil

[7]Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

[8]Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

## ABSTRACT

Numerous genotyping techniques based on different principles and with different costs and levels of resolution are currently available for understanding the transmission dynamics of brucellosis worldwide. We aimed to compare the population structure of the genomes of 53 Brazilian *Brucella abortus* isolates using eight different genotyping methods: multiple-locus variable-number tandem-repeat analysis (MLVA8, MLVA11, MLVA16), multilocus sequence typing (MLST9, MLST21), core genome MLST (cgMLST) and two techniques based on single nucleotide polymorphism (SNP) detection (parSNP and NASP) from whole genomes. The strains were isolated from six different Brazilian states between 1977 and 2008 and had previously been analyzed using MLVA8, MLVA11, and MLVA16. Their whole genomes were sequenced, assembled, and subjected to MSLT9 MLST21, cgMLST, and SNP analyses. All the

genotypes were compared by hierarchical grouping method based on the average distances between the correlation matrices of each technique. MLST9 and MLST21 had the lowest level of resolution, both revealing only four genotypes. MLVA8, MLVA11, and MLVA16 had progressively increasing levels of resolution as more *loci* were analyzed, identifying 6, 16, and 44 genotypes, respectively. cgMLST showed the highest level of resolution, identifying 45 genotypes, followed by the SNP-based methods, both of which had 44 genotypes. In the assessed population, MLVA was more discriminatory than MLST and was easier and cheaper to perform. SNP techniques and cgMLST provided the highest levels of resolution and the results from the two methods were in close agreement. In conclusion, the choice of genotyping technique can strongly affect one's ability to make meaningful epidemiological conclusions but is dependent on available resources: while the VNTR based techniques are more indicated to high prevalence scenarios, the WGS methods are the ones with the best discriminative power and therefore recommended for outbreaks investigation.

**KEY WORDS:** Brucellosis, epidemiology, MLST, MLVA, cgMLST, SNP

## INTRODUCTION

Brucellosis is one of the world's major bacterial zoonoses that affects livestock and humans (WHO, 1997; WHOA, 2018). In cattle, the disease is mainly caused by Brucella abortus (CORBEL; ELBERG; COSIVI, 2006), symptoms include placentitis and abortion, followed by reduction of milk production and secretion of Brucella in milk (BERNUES; MANRIQUE; MAZA, 1997; CARVALHO NETA; MOL; XAVIER; PAIXAO *et al.*, 2010; MCDERMOTT; GRACE; ZINSSTAG, 2013). Bovine brucellosis is endemic and has a great economic impact worldwide; according to World Organization for Animal Health (WOAH) and Food and Agriculture Organization (FAO) data, 200,273 cases of bovine brucellosis caused by *B. abortus* were reported from 2005 to 2019, averaging 13.352 cases/year and incidence of 6.35 cases per 100,000 cattle/year (FAO, 2003; WHOA, 2021). The disease is endemic in Brazil, where the prevalence of positive herds varies between 0.91% (95% confidence interval [CI]: 0.30–2.11%) in Santa Catarina and 30.6% (95% CI: 27.40-34.0%) in Mato Grosso do Sul (FERREIRA NETO; SILVEIRA; ROSA; GONÇALVES *et al.*, 2016).

Due to the economic burden resulting from infection by *B. abortus* and its impact on public health, the control and eradication of brucellosis in cattle have been the goals of various countries (ZHANG; HUANG; WU; LIU *et al.*, 2018), including Brazil. In 2001, Brazil

implemented the Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose Animal (PNCEBT - National Program for Control and Eradication of Animal Brucellosis and Tuberculosis) to reduce the occurrence of the disease (BRASIL, 2001). Measures for controlling bovine brucellosis in the country include the mandatory vaccination of young female calves with S19 or RB51, various hygiene measures such as trade restrictions, test-and-slaughter of seropositive animals, and the voluntary certification of brucellosis-free herds (LAGE; POESTER; GONÇALVES, 2005). The current brucellosis surveillance system in Brazil is primarily based on serological testing and slaughtering of positive animals and, in rare cases, the isolation and biotyping of the pathogen, with a few studies reporting bacterial genotyping by multiple-locus variable number of tandem repeat (VNTR) analysis (MLVA) or multilocus sequence type (MLST) methods (ANDRADE; PEREIRA; SOARES FILHO; SOUZA *et al.*, 2023; DORNELES; SANTANA; ALVES; PAULETTI *et al.*, 2014; MEGID; PAES; LISTONI; RIBEIRO *et al.*, 2005; MINHARRO; SILVA MOL; DORNELES; PAULETTI *et al.*, 2013; OLIVEIRA; DORNELES; SOARES; FONSECA *et al.*, 2017).

The use of genotyping tools for intraspecific classification of *Brucella* spp. has been of great importance for improving epidemiological surveillance (ABDEL-GLIL; THOMAS; BRANDT; MELZER *et al.*, 2022). In fact, molecular epidemiology has become a widely used approach for understanding the evolution of brucellosis and the transmission and movement of strains in countries such as Costa Rica, Egypt, and Italy (GAROFOLO; DI GIANNATALE; PLATONE; ZILLI *et al.*, 2017; KHAN; MELZER; SAYOUR; SHELL *et al.*, 2021; SUÁREZ-ESQUIVEL; HERNÁNDEZ-MORA; RUIZ-VILLALOBOS; BARQUERO-CALVO *et al.*, 2020). Several molecular techniques are currently employed worldwide to elucidate the circulation dynamics of *Brucella* spp. strains, including MLVA and MLST, core-genome multilocus sequence type (cgMLST) analysis and the analysis of single nucleotide polymorphisms (SNPs) obtained from whole-genome sequencing (WGS) (SUÁREZ-ESQUIVEL; CHAVES-OLARTE; MORENO; GUZMÁN-VERRI, 2020). Thus, the simultaneous and systematic implementation of all these techniques in the same group of strains allows not only to compare their different levels of resolution, but also to guide optimal strategies for epidemiological surveillance of brucellosis in Brazil and elsewhere.

MLVA is based on the number of copies of certain repetitive noncoding sequences known as VNTRs (in *Brucella* spp., 8, 11 or 16 loci are traditionally assessed) (AL DAHOUK; LE FLÈCHE; NÖCKLER; JACQUES *et al.*, 2007; LE FLÈCHE; JACQUES; GRAYON; AL DAHOUK *et al.*, 2006) investigated by means of PCR, thus making this a simple and low-cost

approach (LI; RAOULT; FOURNIER, 2009). MLST, on the other hand, uses alleles as the unit of comparison determined from the sequencing of housekeeping genes (generally 9 or 21) (WHATMORE; KOYLASS; MUCHOWSKI; EDWARDS-SMALLBONE *et al.*, 2016; WHATMORE; PERRETT; MACMILLAN, 2007). As this technique detects mutations, insertions, and deletions, sequencing must be performed (typically Sanger) to evaluate each locus (LI; RAOULT; FOURNIER, 2009). The same MLST principle can be applied to the core genome (cgMLST) technique but instead using the gene-by-gene approach, it is based on the comparison of all genes present in the cgMLST of the analyzed group from data generated from WGS (ABDEL-GLIL; THOMAS; BRANDT; MELZER *et al.*, 2022). Similar to cgMLST, the SNP detection technique can be applied to WGS data through the identification of mutations in the sequences. It is a technique with a higher complexity compared to MLST and MLVA and requires expertise in bioinformatics; however, its high level of resolution and discriminatory power stand out as major advantages (SUÁREZ-ESQUIVEL; CHAVES-OLARTE; MORENO; GUZMÁN-VERRI, 2020).

Given that, herein, we compared *B. abortus* population structures based on genotyping methods with different levels of resolution (MLVA8, MLVA11, MLVA16, MLST9, MLST21, cgMLST, parSNP and the Northern Arizona SNP Pipeline [NASP]) by analyzing 53 Brazilian genomes, to provide useful information for bovine brucellosis surveillance programs.

## MATERIALS AND METHODS

### Bacterial strains, WGS, assembly and annotation

Fifty-three *B. abortus* strains were isolated from cattle in six different Brazilian states (Figure 1), from 1977 to 2008, and were identified by microbiological and molecular tests (MINHARRO; SILVA MOL; DORNELES; PAULETTI *et al.*, 2013; PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015). The Illumina HiSeq 2500 platform (Illumina, USA) was used for WGS of the 53 *B. abortus* strains. The quality of the sequencing products was assessed by FASTQC (ANDREWS; KRUEGER; SECONDS-PICHON; BIGGINS *et al.*, 2010). A *de novo* sequence assembly was performed using SPAdes v. 3.13.0 (BANKEVICH; NURK; ANTIPOV; GUREVICH *et al.*, 2012) or Edena v. 3.131028 (BANKEVICH; NURK; ANTIPOV; GUREVICH *et al.*, 2012). The assembly products were ordered in CONTIGuator (GALARDINI; BIONDI; BAZZICALUPO; MENGONI, 2011) using *B. abortus* strain 9-941, A19 or BAB8416 as a reference (HALLING; PETERSON-BURCH; BRICKER; ZUERNER

*et al.*, 2005; LI; KANG; LIN; JIA *et al.*, 2019; WANG; WANG; SUN; BATEER *et al.*, 2020) (Supplementary Table S1). Gaps were closed automatically using the programs GFinisher (GUIZELINI; RAITTZ; CRUZ; SOUZA *et al.*, 2016) and GAPblaster (DE SÁ; MIRANDA; VERAS; DE MELO *et al.*, 2016) and manually using BLAST (https://blast.ncbi.nlm.nih.gov/). The closed genomes were subjected to curatorial analysis and coverage in the CLC (QIAGEN, 2020) and PROKKA was used for genome annotation (SEEMANN, 2014). The genomes were deposited in the National Center for Biotechnology Information (NCBI) database, compiled under a single Bioproject: PRJNA750793. Epidemiological information of the strains can be found in Table 1.

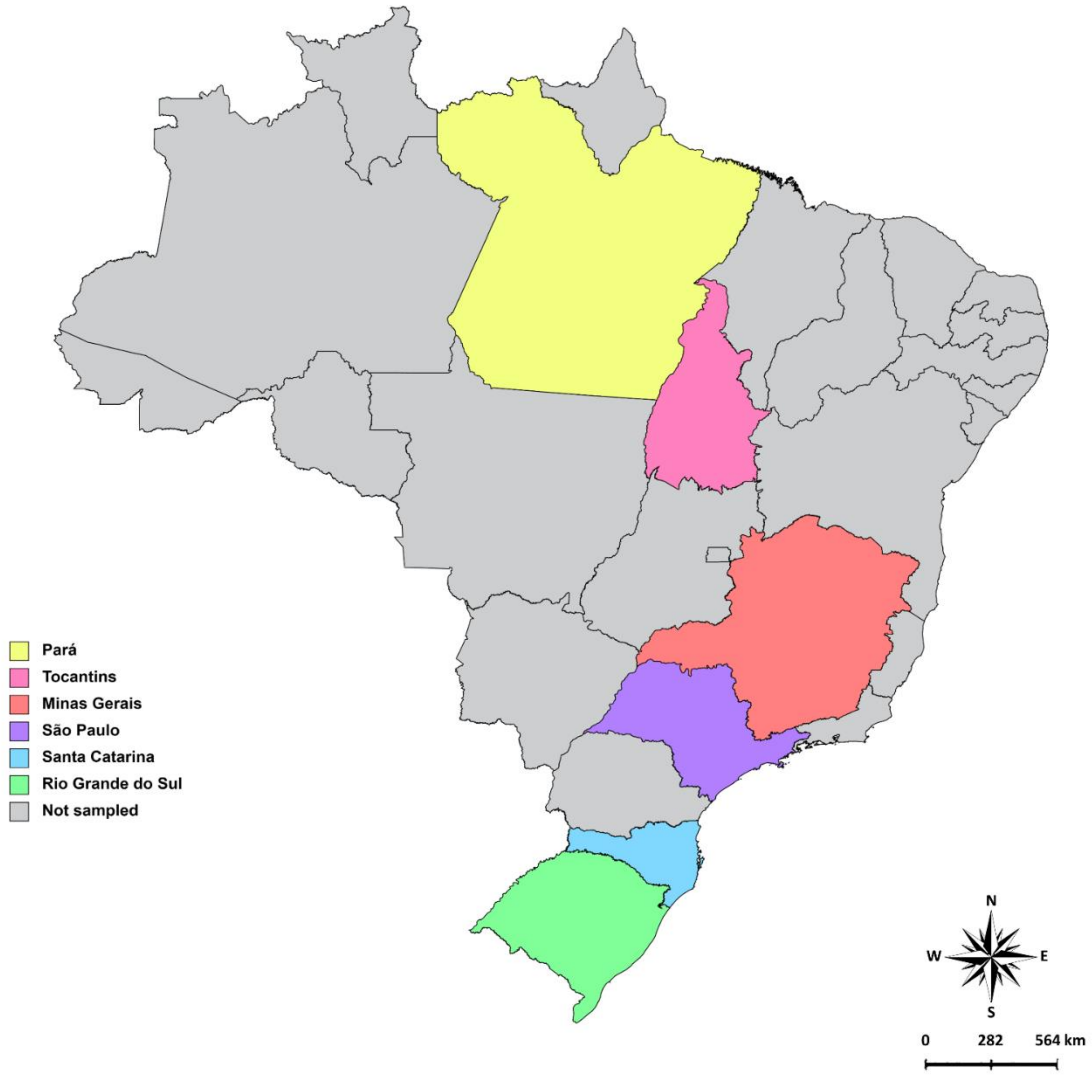


Figure 1: Geographical map of the *Brucella abortus* strains studied, isolated from 1977 to 2008.

Table 1: Epidemiological information of the strains and genotypes under each implemented method.

| Strain | State | Year | Biovar | MLVA | | | MLST | | cgMLST | SNP | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 8 | 11 | 16 | 9 | 21 | | parSNP | NASP |
| LBAB001 | MG | 2006 | 1 | 4 | 10 | 16 | 1 | 1 | 3 | 19 | 19 |
| LBAB003 | MG | 2005 | 6 | 1 | 3 | 7 | 3 | 3 | 36 | 42 | 41 |
| LBAB004 | MG | 2007 | 6 | 1 | 1 | 3 | 3 | 3 | 40 | 40 | 40 |
| LBAB005 | TO | 2007 | 1 | 4 | 10 | 29 | 1 | 1 | 9 | 13 | 13 |
| LBAB007 | PA | 2008 | 3 | 1 | 4 | 8 | 3 | 3 | 43 | 38 | 38 |
| LBAB008 | PA | 2008 | 6 | 1 | 1 | 2 | 3 | 3 | 39 | 39 | 39 |
| LBAB009 | PA | 2008 | 1 | 2 | 5 | 10 | 1 | 1 | 14 | 23 | 23 |
| LBAB010 | PA | 2008 | 1 | 4 | 10 | 15 | 1 | 1 | 10 | 16 | 16 |
| LBAB011 | TO | 2008 | 3 | 1 | 1 | 1 | 3 | 3 | 42 | 44 | 44 |
| LBAB012 | PA | 2008 | 3 | 1 | 1 | 4 | 3 | 3 | 41 | 37 | 37 |
| LBAB013 | PA | 2008 | 1 | 4 | 10 | 27 | 1 | 1 | 26 | 22 | 22 |
| LBAB014 | MG | - | 1 | 4 | 14 | 41 | 1 | 1 | 16 | 27 | 27 |
| LBAB015 | MG | - | 1 | 4 | 12 | 37 | 1 | 1 | 17 | 26 | 26 |
| LBAB016 | MG | - | 1 | 4 | 12 | 36 | 1 | 1 | 15 | 26 | 26 |
| LBAB017 | MG | - | 1 | 4 | 12 | 36 | 1 | 1 | 17 | 26 | 26 |
| LBAB018 | MG | - | 1 | 4 | 11 | 30 | 1 | 1 | 18 | 25 | 25 |
| LBAB019 | RS | 2006 | 1 | 4 | 10 | 23 | 1 | 1 | 21 | 4 | 4 |
| LBAB020 | RS | 2006 | 1 | 4 | 10 | 24 | 1 | 1 | 20 | 2 | 2 |
| LBAB021 | SC | 2007 | 1 | 4 | 10 | 22 | 1 | 1 | 35 | 32 | 32 |
| LBAB022 | RS | 2006 | 1 | 5 | 15 | 42 | 4 | 4 | 4 | 21 | 21 |
| LBAB023 | RS | 2002 | 1 | 4 | 12 | 39 | 3 | 3 | 37 | 41 | 42 |
| LBAB024 | RS | 2002 | 1 | 4 | 10 | 25 | 1 | 1 | 2 | 18 | 18 |
| LBAB025 | RS | 2003 | 1 | 4 | 10 | 24 | 1 | 1 | 19 | 9 | 9 |
| LBAB026 | RS | 2003 | 1 | 4 | 11 | 33 | 1 | 1 | 7 | 14 | 14 |
| LBAB027 | SP | - | 1 | 4 | 11 | 34 | 1 | 1 | 22 | 3 | 3 |
| LBAB028 | SP | - | 6 | 1 | 2 | 5 | 3 | 3 | 38 | 43 | 43 |
| LBAB029 | SP | 2008 | 6 | 1 | 2 | 6 | 3 | 3 | 37 | 41 | 42 |
| LBAB030 | MG | 2006 | 1 | 4 | 10 | 16 | 1 | 1 | 28 | 8 | 8 |
| LBAB031 | MG | 2006 | 1 | 4 | 10 | 17 | 1 | 1 | 27 | 7 | 7 |
| LBAB032 | RS | 1977 | 1 | 4 | 8 | 13 | 1 | 1 | 29 | 29 | 29 |
| LBAB033 | SP | - | 6 | 2 | 6 | 11 | 1 | 1 | 20 | 2 | 2 |
| LBAB034 | SC | 2007 | 1 | 4 | 10 | 21 | 1 | 1 | 35 | 32 | 32 |
| LBAB035 | RS | 2004 | 1 | 4 | 10 | 28 | 1 | 1 | 23 | 5 | 5 |
| LBAB036 | RS | 2004 | 1 | 4 | 11 | 34 | 1 | 1 | 23 | 5 | 5 |
| LBAB037 | RS | 1996 | 1 | 4 | 10 | 26 | 1 | 1 | 8 | 12 | 12 |
| LBAB038 | RS | 1977 | 1 | 4 | 10 | 18 | 2 | 2 | 45 | 28 | 28 |
| LBAB039 | SC | - | 1 | 4 | 10 | 19 | 1 | 1 | 33 | 34 | 34 |
| LBAB040 | SC | 2007 | 1 | 4 | 10 | 19 | 1 | 1 | 32 | 31 | 31 |
| LBAB041 | SC | - | 1 | 4 | 10 | 19 | 1 | 1 | 34 | 33 | 33 |
| LBAB042 | RS | - | 1 | 4 | 11 | 35 | 1 | 1 | 25 | 1 | 1 |
| LBAB043 | RS | 2004 | 1 | 2 | 5 | 9 | 1 | 1 | 24 | 6 | 6 |
| LBAB044 | RS | 2006 | 1 | 4 | 11 | 32 | 4 | 4 | 5 | 20 | 20 |
| LBAB045 | RS | 2007 | 1 | 4 | 11 | 31 | 1 | 1 | 1 | 17 | 17 |
| LBAB046 | RS | - | 1 | 4 | 12 | 38 | 1 | 1 | 20 | 2 | 2 |
| LBAB047 | SC | - | 1 | 4 | 10 | 21 | 1 | 1 | 35 | 32 | 32 |
| LBAB048 | SP | - | 1 | 4 | 13 | 40 | 1 | 1 | 6 | 10 | 10 |
| LBAB049 | SP | - | 1 | 6 | 16 | 43 | 1 | 1 | 13 | 11 | 11 |
| LBAB050 | SP | - | 1 | 4 | 10 | 20 | 1 | 1 | 12 | 24 | 24 |
| LBAB051 | SP | - | 1 | 4 | 9 | 14 | 1 | 1 | 44 | 35 | 35 |
| LBAB052 | SP | - | 1 | 4 | 10 | 19 | 1 | 1 | 31 | 30 | 30 |
| LBAB053 | SP | - | 1 | 4 | 10 | 20 | 1 | 1 | 11 | 15 | 15 |
| LBAB054 | SP | - | 1 | 6 | 16 | 44 | 1 | 1 | 12 | 24 | 24 |
| LBAB055 | RS | 2002 | 4 | 3 | 7 | 12 | 1 | 1 | 30 | 36 | 36 |

MG: Minas Gerais; TO: Tocantins; PA: Pará; RS: Rio Grande do Sul; SC: Santa Catarina; SP: São Paulo; -: Unknown

### Genotyping methods

Analyses of MLVA8, MLVA11 and MLVA16 were previously performed by MINHARRO; SILVA MOL; DORNELES; PAULETTI *et al.* (2013). MLST9 and MLST21 analyses were performed using the MLST program (SEEMANN, 2022; WHATMORE; KOYLASS; MUCHOWSKI; EDWARDS-SMALLBONE *et al.*, 2016; WHATMORE; PERRETT; MACMILLAN, 2007) and the PubMLST platform (https://pubmlst.org/). cgMLST was implemented using the chewBBACA pipeline (SILVA; MACHADO; SILVA; ROSSI *et al.*, 2017), and the WGS-SNP based approaches were performed in parSNP and NASP (SAHL; LEMMER; TRAVIS; SCHUPP *et al.*, 2016; TREANGEN; ONDOV; KOREN; PHILLIPPY, 2014).

### Statistical analysis

The Hunter-Gaston discriminatory index (HGDI) and its 95% confidence interval (CI) were calculated to the MLVA8, MLVA11, MLVA16, MLST9, MLST 21, cgMLST, parSNP and NASP genotyping results using Simpson's diversity index (SDI) and the adjusted Wallace (AW) test using an online tool (http://www.comparingpartitions.info/?link=Tool) (HUNTER; GASTON, 1988; SEVERIANO; PINTO; RAMIREZ; CARRIÇO, 2011). Comparisons of agreement between the approaches were performed using a hierarchical grouping method based on the average distances between the correlation matrices of each technique using the "dendextend" package of R version 4.2.1 (GALILI, 2015; R, 2022). The trees were visualized using iTOL (https://itol.embl.de/).

## RESULTS

### Genomes

The 53 genomes were deposited with no gaps in the format of two chromosomes, with an average guanine–cytosine (GC) content of 57.25%, an average size of 3,283,155 base pairs and an average coverage of 264x. The average number of coding sequences (CDS) was 3,124 ± 9, ranging from 3,101 to 3,138.

### Genotyping methods

In Figure 2, MLVA8, MLVA11 and MLVA16 identified 6, 16 and 44 different genotypes, respectively. In these three methods, biovars 3 and 6 grouped together genotypically. Likewise, both MLST9 and MLST21, which identified only four genotypes, also indicated the separation

of the isolates according biovars in the corresponding phylograms. However, despite the differentiation of only few genotypes among the studied genomes, two new alleles at the *gap* and *mutL* loci were found and deposited on the PubMLST (https://pubmlst.org/). The cgMLST analysis identified 45 different genotypes through the analysis of 2,829 genes. In this technique, as well as the SNP-based techniques (parSNP with 1,973 and NASP with 1,987 polymorphisms), it was possible to observe not only the separation of biovars 3 and 6 strains (except for one biovar 6 isolate) but also a clearer division of the isolates from Santa Catarina, the only Brazilian State where the isolates exhibited clustering (indicated in blue in Figure 1).

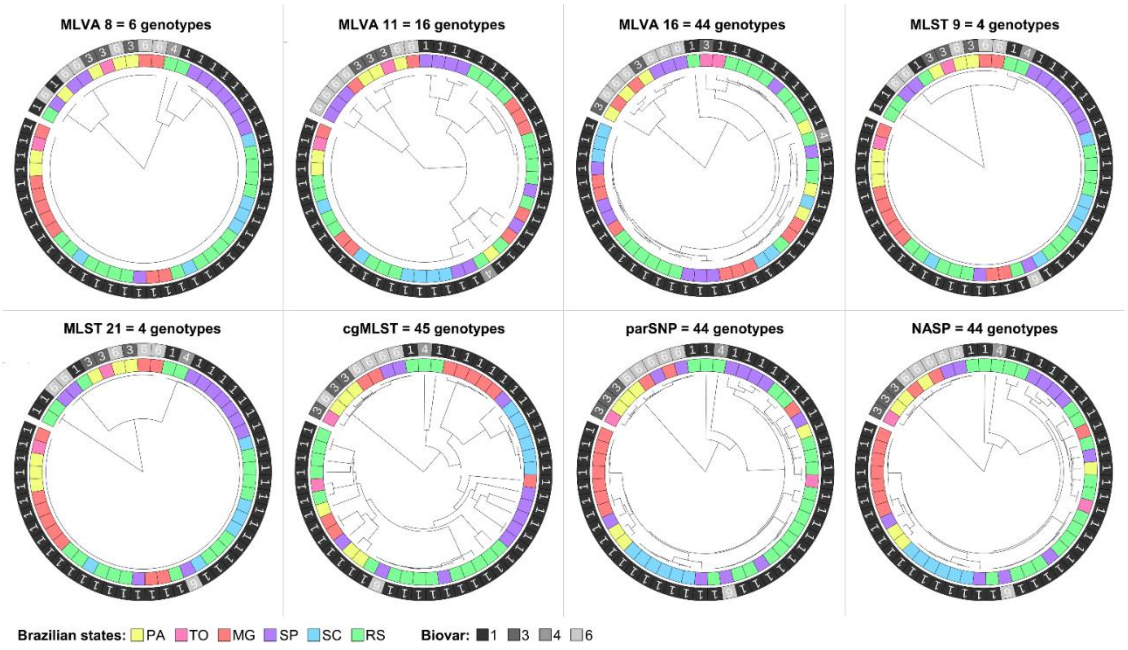


Figure 2: Dendrograms generated for each of the genotyping techniques applied to the 53 *Brucella abortus* strains isolated from 1977 to 2008. The Brazilian states are represented in the inner circles, with the colors corresponding to the states from the map in Figure 1, and biovars are represented in the outer circle in grayscale. MG: Minas Gerais; TO: Tocantins; PA: Pará; RS: Rio Grande do Sul; SC: Santa Catarina; SP: São Paulo.

The genotyping technique showing the highest discrimination between genomes was cgMLST, followed by the WGS-SNP based approaches (parSNP and NASP had the same level of resolution), MVLA16, MLVA11, MLVA8 and finally MLST9 and MLST21, these last two showed the same level of resolution. The HGDI values and their 95% confidence intervals for each genotyping method are shown in Table 2.

Table 2: Diversity index (HGDI) calculated for each technique (MLVA8, MLVA11, MLVA16, MLST9, MLST21, cgMLST, parSNP and NASP) from the genomes of *B. abortus* isolated from cattle, Brazil, from 1977 to 2008.

| Technique | Genotypes number | HGDI | 95% Confidence interval |
|---|---|---|---|
| MLVA8 | 6 | 0.451 | (0.292-0.610) |
| MLVA11 | 16 | 0.811 | (0.718-0.905) |
| MLVA16 | 44 | 0.992 | (0.983-1.000) |
| MLST9 | 4 | 0.384 | (0.235-0.532) |
| MLST21 | 4 | 0.384 | (0.235-0.532) |
| cgMLST | 45 | 0.992 | (0.985-1.000) |
| parSNP | 44 | 0.991 | (0.983-0.999) |
| NASP | 44 | 0.991 | (0.983-0.999) |

Hierarchical grouping based on the average distance between the correlation matrices of each technique depicted a strong agreement between MLST9 and MLST21, among MLVA8, MLVA11 and MLVA16 and among the WGS-based techniques (cgmLST, parSNP and NASP) (Figure 3). A phylogram comparing all the techniques against parSNP can be seen in Supplementary Files S1-S7.

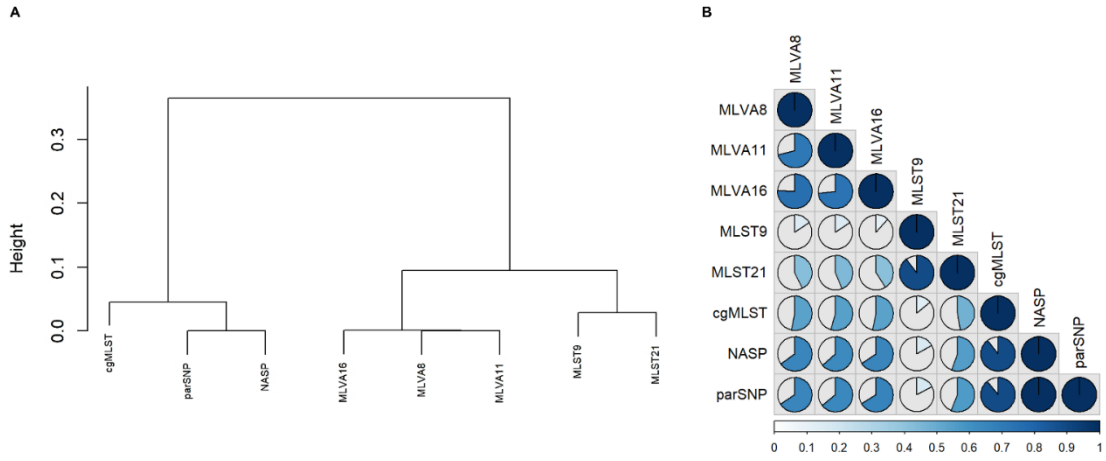


Figure 3: Hierarchical grouping by the mean distance between the correlation matrices of each technique using the 53 Brazilian genomes of *Brucella abortus* strains isolated from 1977 to 2008.

## DISCUSSION

In this study, we compared the performance of eight genotyping methods to assess the phylogenetic relationships among *B. abortus* isolates from different Brazilian states with the aim to determine which would be the most suitable techniques for an epidemiological investigation in different contexts of brucellosis prevalence, taking also into account the different levels of work, infrastructure required and the overall costs of the different analysis methods.

All techniques produced results that discriminated most of the genomes splitting them into a group composed of biovars 3 and 6 (with few exceptions), and another composed of biovars 1 and 4. Techniques based on SNPs (NASP and parSNP) outstand for being able to separate biovars 3 and 6 into distinct clades. However, it should be noted that this biovar grouping observed at different levels of resolution by all techniques possibly occurred due to the epidemiological link present among the strains and probably does not reflect the presence of any VNTRs or specific mutations of a given biovar in the genomic regions analyzed. In fact, the exception observed in a biovar 6 genome, apart from the others, supports discussions pointed out by a recent literature review that the biovar classification that takes into account phenotypic characteristics does not necessarily have concordance with the observed genotype (WHATMORE; FOSTER, 2021).

MLVA16 and WGS-based methods (cgMLST, parSNP and NASP) showed the greatest discriminatory power, with Simpson's ID values very similar and greater than 0.990, as also reported for *B. melitensis* by JANOWICZ; DE MASSIS; ANCORA; CAMMÀ *et al.* (2018). Nonetheless, despite showing very similar HGDI values (MLVA16 vs. WGS-based methods), WGS-based methods have a greater resolving power than MLVA16 for closely related strains (Figure 2). Since WGS-based techniques cover almost the entire genome, as they consider the entire core genome (extremely conserved among *B. abortus* strains) and therefore provide a high level of detail, it is ideal for differentiating closely related strains. MLVA-based techniques, in contrast, are based on the evaluation of a limited portion of the bacterial genome, depending on the number of repeats of the analyzed VNTRs. However, the great advantage of these techniques is that, despite covering only approximately 0.1% of the genome (WHATMORE; FOSTER, 2021), in endemic areas they can nonetheless identify close relationships among strains. Thus, in epidemiological scenarios in which bovine brucellosis still occurs, as found in most states of Brazil (FERREIRA NETO; SILVEIRA; ROSA; GONÇALVES *et al.*, 2016), MLVA16 is a technique that can satisfactorily differentiate unrelated isolates, with a much lower complexity of execution and cost than analyses based on WGS (LI; RAOULT; FOURNIER, 2009). Moreover, the MLVA database contains thousands of genotypes, allowing for a comprehensive dataset to place strains into a global and sometimes local context. Similarly, as some laboratories in Latin America, Eastern Europe, and Africa lack the facilities to carry out WGS and have limited resources to cover the logistics of sending samples for sequencing, MLVA16 can be a very useful epidemiological tool in the surveillance of *Brucella* spp., optimizing disease control and prevention actions in these regions (LEDWABA; GOMO;

LEKOTA; LE FLÈCHE *et al.*, 2019). It is also worth mentioning that MLVA16 in these contexts has the great advantage of a shorter execution time, since the WGS requires not only bioinformatic skills but also computational power for data processing (SUÁREZ-ESQUIVEL; CHAVES-OLARTE; MORENO; GUZMÁN-VERRI, 2020).

On the other hand, in scenarios in which brucellosis is close to eradication, where the herd prevalence is below 1%, as found in the state of Santa Catarina (BAUMGARTEN; VELOSO; GRISI-FILHO; FERREIRA *et al.*, 2016), Brazil, and in several other parts of the world, cgMLST and the SNP-based techniques had high discriminatory power and grouped the genomes with greater confidence [HGDI = 0.992 and 95% (CI) = 0.985–1.000], besides it showed epidemiological concordance. This can be explained by the fact that point mutations and deletions are evolutionarily more stable than VNTRs and that this strategy takes into account thousands of genes and SNPs distributed throughout the genome, and consequently has a higher resolution (SUN; JING; DI; YAN *et al.*, 2017; WHATMORE; FOSTER, 2021). Furthermore, it is worth mentioning that for developed countries, the disadvantages mentioned above for techniques based on WGS are likely not an impairment to its implementation, being always necessary to evaluate which technique is most suitable for each situation, considering not only technical criteria.

In addition, some aspects of the WGS-based techniques should be highlighted: we observed slightly greater diversity under cgMLST (45 genotypes) using a gene-by-gene approach than under SNP-based methods evaluating polymorphisms (44 genotypes). In cgMLST, each allelic change is counted as a single event, regardless of the number of nucleotide polymorphisms involved (SUÁREZ-ESQUIVEL; CHAVES-OLARTE; MORENO; GUZMÁN-VERRI, 2020), allowing the combination of several SNPs in the same gene to generate a greater number of alleles compared to the counting of mutations under the parSNP and NASP methods (SAHL; LEMMER; TRAVIS; SCHUPP *et al.*, 2016; TREANGEN; ONDOV; KOREN; PHILLIPPY, 2014). Nevertheless, this difference may not be observed in other studies, since cgMLST is a WGS genotyping system based on a predefined set of core genes (ABDEL-GLIL; THOMAS; BRANDT; MELZER *et al.*, 2022) that may vary according to the group of genomes being analyzed.

The two applied SNP-based techniques showed very high agreement, as expected. ParSNP is currently one of the tools most commonly used for this purpose; however, it is limited to only using assembled genomes as input files. NASP in this aspect, despite being less widespread, is much more flexible in terms of data types, allowing the use of read files (SAHL;

LEMMER; TRAVIS; SCHUPP *et al.*, 2016). Therefore, the application of NASP can be extremely useful in outbreak investigations since it does not require the performance of the entire genome assembly process for the subsequent analysis of polymorphisms, thereby reducing the time needed to obtain the results and implement the appropriate measures based on their interpretation.

Conversely, the MLST results showed very low discrimination among the isolates, confirmed by the low HGDI values obtained when analyzing either 9 or 21 loci (HGDI = 0.384). This occurs mainly due to the lack of diversity within the genus *Brucella* at the investigated loci, composed of only 9 or 21 housekeeping genes, which makes the analysis based on only a few regions that are highly conserved in this genus. In fact, it is worth mentioning that the typical *Brucella* genomes show more than 90% genomic similarity (RAJENDHRAN, 2021), and considering only housekeeping genes this level of conservation is even greater, limiting the resolving power for this genus (WHATMORE; FOSTER, 2021). In this context, it is important to consider that if this had been the only typing technique used in the present study, it would not be useful in providing information on the relatedness of the strains from the epidemiological point of view. Thus, considering that the performance of this technique is slightly more complex than MLVA and generates less-informative results in distinguishing unrelated isolates, it is not recommended as a first choice for epidemiological surveillance due to its low cost-benefit ratio.

In conclusion, we observed that different techniques are more suitable depending on the available time, resources, specialized bioinformatic skills and epidemiological scenario (i.e. the relatedness between strains). In the investigated context, MLVA16 was a simpler and more ideal technique for countries that are still in the disease-control phase, while the SNP and cgMLST techniques are better choices for outbreak investigations and for surveillance in countries in the eradication phase or where brucellosis has already been eradicated.

## SUPPLEMENTARY FILES

https://drive.google.com/drive/folders/1vOtahFFRv7pk0kWVbLRhQkvFIDeOEzKy?usp=drive_link

S1 – S7: Supplementary files with the dendrograms comparing MLVA8 (S1), MLVA11 (S2), MLVA16 (S3), MLST9 (S4), MLST21 (S5), cgMLST (S6) and NASP (S7) against the parSNP technique in analyzing 53 genomes of Brucella abortus strains isolated in Brazil between 1977 and 2008.

**ACKNOWLEDGMENTS**

This study was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq [423557/2016-5], Fundação de Amparo à Pesquisa do Estado de Minas Gerais – Fapemig [APQ-00283/2016 and RED-00132-22], Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes and Network of Omics Sciences (RECOM). DO'C received a travel grant from CAPES/PrInt/UFLA [52/2019]. CRP received a scholarship from CAPES/PrInt/UFLA [01/2020]. EMSD is also grateful to CNPq for her fellowship.

## REFERENCES

ABDEL-GLIL, M. Y.; THOMAS, P.; BRANDT, C.; MELZER, F. *et al.* Core genome multilocus sequence typing scheme for improved characterization and epidemiological surveillance of pathogenic *Brucella*. **Journal of Clinical Microbiology**, 60, n. 8, p. e0031122, Aug 17 2022.

AL DAHOUK, S.; LE FLÈCHE, P.; NÖCKLER, K.; JACQUES, I. *et al.* Evaluation of *Brucella* MLVA typing for human brucellosis. **Journal of Microbiological Methods**, 69, n. 1, p. 137-145, 2007.

ANDRADE, R. S.; PEREIRA, C. R.; SOARES FILHO, P. M.; SOUZA, P. G. *et al.* Phenotypic and genotypic characterization of *Brucella abortus* biovar 4 isolates from cattle in Brazil. **Ciência Rural**, 53, n. 6, p. e20210564, 2023.

ANDREWS, S.; KRUEGER, F.; SECONDS-PICHON, A.; BIGGINS, F. *et al.* **A quality control tool for high throughput sequence data.** Babraham Bioinformatics, 2010.

BANKEVICH, A.; NURK, S.; ANTIPOV, D.; GUREVICH, A. A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **Journal of Computational Biology**, 19, n. 5, p. 455-477, 2012.

BAUMGARTEN, K. D.; VELOSO, F. P.; GRISI-FILHO, J. H. H.; FERREIRA, F. *et al.* Prevalence and risk factors for bovine brucellosis in the State of Santa Catarina, Brazil. **Semina: Ciências Agrárias**, 37, n. 5, p. 3425-3435, 2016.

BERNUES, A.; MANRIQUE, E.; MAZA, M. T. Economic evaluation of bovine brucellosis and tuberculosis eradication programmes in a mountain area of Spain. **Preventive Veterinary Medicine**, 30, n. 2, p. 137-149, 1997.

BRASIL. Instituir o Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose Animal. Instrução Normativa Ministerial n. 02 de 10 de janeiro de 2001. MINISTÉRIO DA AGRICULTURA, P. e. A. Brasília: Diário Oficial da União. 2 2001.

CARVALHO NETA, A. V.; MOL, J. P.; XAVIER, M. N.; PAIXAO, T. A. *et al.* Pathogenesis of bovine brucellosis. **The Veterinary Journal**, 184, n. 2, p. 146-155, May 2010.

CORBEL, M. J.; ELBERG, S. S.; COSIVI, O. **Brucellosis in humans and animals**. Geneva: World Health Organization, 2006. 189 p.

DE SÁ, P. H. C. G.; MIRANDA, F.; VERAS, A.; DE MELO, D. M. *et al.* GapBlaster—a graphical gap filler for prokaryote genomes. **PLoS One**, 11, n. 5, p. e0155327, 2016.

DORNELES, E. M. S.; SANTANA, J. A.; ALVES, T. M.; PAULETTI, R. B. *et al.* Genetic stability of *Brucella abortus* isolates from an outbreak by multiple-locus variable-number tandem repeat analysis (MLVA16). **BMC Microbiology**, 14, n. 1, p. 186, 2014/07/11 2014.

FAO. World Agriculture: Towards 2015/2030 - A FAO Perspective (Food and Agriculture Organization of the United Nations), Livestock production: production and productivity. : Food and Agriculture Organization 2003.

FERREIRA NETO, J. S.; SILVEIRA, G. B.; ROSA, B. M.; GONÇALVES, V. S. P. *et al.* Analysis of 15 years of the National Program for the Control and Eradication of Animal Brucellosis and Tuberculosis, Brazil. **Semina: Ciências Agrárias**, 37, n. 5, suppl.2, p. 3385-3402, 2016.

GALARDINI, M.; BIONDI, E. G.; BAZZICALUPO, M.; MENGONI, A. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. **Source code for biology and medicine**, 6, n. 1, p. 11, 2011.

GALILI, T. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. **Bioinformatics**, 31, n. 22, p. 3718-3720, 2015.

GAROFOLO, G.; DI GIANNATALE, E.; PLATONE, I.; ZILLI, K. *et al.* Origins and global context of *Brucella abortus* in Italy. **BMC Microbiology**, 17, n. 1, p. 28, 2017/02/02 2017.

GUIZELINI, D.; RAITTZ, R. T.; CRUZ, L. M.; SOUZA, E. l. M. *et al.* GFinisher: a new strategy to refine and finish bacterial genome assemblies. **Scientific reports**, 6, p. 34963, 2016.

HALLING, S. M.; PETERSON-BURCH, B. D.; BRICKER, B. J.; ZUERNER, R. L. *et al.* Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. **Journal of Bacteriology**, 187, n. 8, p. 2715-2726, 2005.

HUNTER, P. R.; GASTON, M. A. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. **Journal of Clinical Microbiology**, 26, n. 11, p. 2465-2466, 1988.

JANOWICZ, A.; DE MASSIS, F.; ANCORA, M.; CAMMÀ, C. *et al.* Core genome multilocus sequence typing and single nucleotide polymorphism analysis in the epidemiology of *Brucella melitensis* infections. **Journal of Clinical Microbiology**, 56, n. 9, p. e00517-00518, 2018.

KHAN, A. U.; MELZER, F.; SAYOUR, A. E.; SHELL, W. S. *et al.* Whole-Genome Sequencing for Tracing the Genetic Diversity of *Brucella abortus* and *Brucella melitensis* Isolated from Livestock in Egypt. **Pathogens**, 10, n. 6, p. 20, Jun-%J-Pathogens 2021.

LAGE, A. P.; POESTER, F. P.; GONÇALVES, V. S. P. Controle da brucelose bovina. **Cadernos Técnicos de Veterinária e Zootecnia** 47, p. 30-41, 2005.

LE FLÈCHE, P.; JACQUES, I.; GRAYON, M.; AL DAHOUK, S. *et al.* Evaluation and selection of tandem repeat loci for a *Brucella* MLVA typing assay. **BMC Microbiology**, 6, n. 1, p. 1-14, 2006.

LEDWABA, M. B.; GOMO, C.; LEKOTA, K. E.; LE FLÈCHE, P. *et al.* Molecular characterization of *Brucella* species from Zimbabwe. **PLoS neglected tropical diseases**, 13, n. 5, p. e0007311, 2019.

LI, W.; RAOULT, D.; FOURNIER, P. E. Bacterial strain typing in the genomic era. **FEMS Microbiology Reviews**, 33, n. 5, p. 892-916, 2009.

LI, X.; KANG, Y.; LIN, L.; JIA, E. *et al.* Genomic Characterization provides new insights for detailed phage-resistant mechanism for *Brucella abortus*. **Frontiers in Microbiology**, 10, p. 917, 2019.

MCDERMOTT, J.; GRACE, D.; ZINSSTAG, J. Economics of brucellosis impact and control in low-income countries. **Revue scientifique et technique**, 32, n. 1, p. 249-261, 2013.

MEGID, J.; PAES, A. C.; LISTONI, F. P.; RIBEIRO, M. G. *et al.* Isolation of *Brucella abortus* from cattle and water buffalo in Brazil. **The Veterinary Record**, 156, n. 5, p. 147, 2005.

MINHARRO, S.; SILVA MOL, J. P.; DORNELES, E. M. S.; PAULETTI, R. B. *et al.* Biotyping and genotyping (MLVA16) of *Brucella abortus* isolated from cattle in Brazil, 1977 to 2008. **Plos One**, 8, n. 12, p. e81152, 2013.

OLIVEIRA, M. S.; DORNELES, E. M. S.; SOARES, P. M. F.; FONSECA, A. A. *et al.* Molecular epidemiology of *Brucella abortus* isolated from cattle in Brazil, 2009–2013. **Acta Tropica**, 166, p. 106-113, 2017/02/01/ 2017.

PAULETTI, R. B.; STYNEN, A. P. R.; MOL, J. P. S.; DORNELES, E. M. S. *et al.* Reduced Susceptibility to Rifampicin and Resistance to Multiple Antimicrobial Agents among Brucella abortus Isolates from Cattle in Brazil. **PLoS One**, 10, n. 7, p. e0132532, 2015.

QIAGEN. **CLC Genomics Workbench**. Versão 20.0. 2020.

R. **R: A language and environment for statistical computing**. Versão 4.2.2. Vienna, Austria: R Core Team, 2022.

RAJENDHRAN, J. Genomic insights into *Brucella*. **Infection, Genetics and Evolution**, 87, p. 104635, 2021/01/01/ 2021.

SAHL, J. W.; LEMMER, D.; TRAVIS, J.; SCHUPP, J. M. *et al.* NASP: an accurate, rapid method for the identification of SNPs in WGS datasets that supports flexible input and output formats. **Microbial Genomics**, 2, n. 8, 2016.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, 30, n. 14, p. 2068-2069, 2014.

SEEMANN, T. **MLST**. Versão 2.22.1. 2022.

SEVERIANO, A.; PINTO, F. R.; RAMIREZ, M.; CARRIÇO, J. A. Adjusted Wallace coefficient as a measure of congruence between typing methods. **Journal of Clinical Microbiology**, 49, n. 11, p. 3997-4000, 2011.

SILVA, M.; MACHADO, M. P.; SILVA, D. N.; ROSSI, M. *et al.* chewBBACA: A complete suite for gene-by-gene schema creation and strain identification. **Microbial Genomics**, 4, 2017.

SUÁREZ-ESQUIVEL, M.; CHAVES-OLARTE, E.; MORENO, E.; GUZMÁN-VERRI, C. *Brucella* genomics: macro and micro evolution. **International Journal of Molecular Sciences**, 21, n. 20, p. 7749, 2020.

SUÁREZ-ESQUIVEL, M.; HERNÁNDEZ-MORA, G.; RUIZ-VILLALOBOS, N.; BARQUERO-CALVO, E. *et al.* Persistence of *Brucella abortus* lineages revealed by genomic characterization and phylodynamic analysis. **PLoS neglected tropical diseases**, 14, n. 4, p. e0008235, 2020.

SUN, M.; JING, Z.; DI, D.; YAN, H. *et al.* Multiple locus variable-number tandem-repeat and single-nucleotide polymorphism-based *Brucella* typing reveals multiple lineages in *Brucella melitensis* currently endemic in China. **Frontiers in Veterinary Science**, 4, p. 215, 2017.

TREANGEN, T. J.; ONDOV, B. D.; KOREN, S.; PHILLIPPY, A. M. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. **Genome Biology**, 15, n. 11, p. 1-15, 2014.

WANG, S.; WANG, W.; SUN, K.; BATEER, H. *et al.* Comparative genomic analysis between newly sequenced *Brucella abortus* vaccine strain A19 and another Brucella abortus vaccine S19. **Genomics**, 112, n. 2, p. 1444-1453, 2020.

WHATMORE, A. M.; FOSTER, J. T. Emerging diversity and ongoing expansion of the genus *Brucella*. **Infection, Genetics and Evolution**, 92, p. 104865, 2021.

WHATMORE, A. M.; KOYLASS, M. S.; MUCHOWSKI, J.; EDWARDS-SMALLBONE, J. *et al.* Extended multilocus sequence analysis to describe the global population structure of the genus *Brucella*: phylogeography and relationship to biovars. **Frontiers in Microbiology**, 7, 2016-December-21 2016. Original Research.

WHATMORE, A. M.; PERRETT, L. L.; MACMILLAN, A. P. Characterisation of the genetic diversity of *Brucella* by multilocus sequencing. **BMC Microbiology**, 7, n. 1, p. 34, 2007/04/20 2007.

WHO. The development of new/improved brucellosis vaccines: Report of a WHO Meeting. Geneva: World Health Organization**:** 48 p. 1997.

WHOA. Brucellosis (*Brucella abortus*, *B. melitensis* and *B. suis*). *In*: HEALTH, W. O. f. A. (Ed.). **Manual of diagnostic tests and vaccines for terrestrial animals 2018**. 7 ed. Paris: World Organisation for Animal Health, 2018. p. 1404.

WHOA. World Animal Health Information System (WAHIS). World Organisation for Animal Health 2021.

ZHANG, N.; HUANG, D.; WU, W.; LIU, J. *et al.* Animal brucellosis control or eradication programs worldwide: a systematic review of experiences and lessons learned. **Preventive Veterinary Medicine**, 160, p. 105-115, 2018.

# CHAPTER 4 - *Brucella abortus* genome quality: implications for the microbial pan-genome

Carine Rodrigues Pereira[1]; Laís Gabrielle Alvarenga Assis[1]; Andrei Giacchetto Felice[2]; Jeffrey T. Foster[3]; David O'Callaghan[4]; Siomar de Castro Soares[2]; Vasco Ariston de Carvalho Azevedo[5]; Andrey Pereira Lage[6]; Elaine Maria Seles Dorneles[1]*.

[1]Faculdade de Zootecnia e Medicina Veterinária, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil

[2]Instituto de Ciências Biológicas e Naturais, Universidade Federal do Triângulo Mineiro, Uberaba, Minas Gerais, Brazil

[3]Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, Arizona, United States

[4]Bacterial Virulence and Infectious Disease, University of Montpellier, Nimes, France

[5]Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

[6]Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

*Address correspondence to author: +55 (35) 3829-1712 – elaine.dorneles@ufla.br – Universidade Federal de Lavras Caixa Postal 3037 – CEP: 37200-900 – Lavras – MG – Brazil

## ABSTRACT:

The study of the genomes of bacterial pathogens has contributed enormously to the knowledge of microorganisms and, therefore, to the control and prevention of zoonoses, such as brucellosis. This study aimed to perform and compare the pan-genomic analysis of 53 strains of *Brucella abortus* isolated from cattle in Brazil, in combination with genomes available on the platform of the National Center for Biotechnology Information (NCBI), before and after a quality curation of the assemblies, to verify the impact that the absence of basic quality parameters may have in the study of the genetic repertoire of the species. The core-genome, accessory genome and singletons were predicted in the pan-genome analyzes using Roary before and after applying the quality filter (309 and 252 genomes respectively). We identified 4,955 genes in the pan-genome after que quality filter, being 4,921 of them present in the NCBI

strains, and 3,361 in the Brazilian strains. The quality filter caused a three-fold increase in the number of genes common to all isolates and reduced the number of singletons by almost seven times. The 53 genomes of *B. abortus* contributed to increase knowledge about the genetic repertoire of the species and have the potential to serve as a basis for future functional studies of this important zoonotic pathogen. In conclusion, the quality curation of assemblies before the performance of a pan-genome analysis proved to be an essential step for an accurate characterization of the genetic repertoire of *B. abortus*. Additionally, the inclusion of the 53 complete genomes from strains isolated in Brazil was responsible for adding 34 genes to the known genetic repertoire of *B. abortus*.

**KEYWORDS:** Brucellosis, whole genome sequence, core-genome, shared-genome, singletons

## INTRODUCTION

Brucellosis is an important bacterial zoonosis associated with reproductive failure in domestic animals and debilitating febrile illness in humans (PAPPAS; PAPADIMITRIOU; AKRITIDIS; CHRISTOU *et al.*, 2006). Although the *Brucella* genus includes 13 species (in current expansion), most infections in cattle are caused by *Brucella abortus* (ABOUT; PASTRE; BOUTROU; MARTINEZ *et al.*, 2023; WHATMORE; FOSTER; EVOLUTION, 2021). Bovine brucellosis is endemic and has large economic impact in many developing regions, such as Africa, Asia, Latin America and Middle East (FRANC; KRECEK; HÄSLER; ARENAS-GAMBOA, 2018; KIIZA; DENAGAMAGE; SERRA; MAUNSELL *et al.*, 2023). The infections caused by *B. abortus* are responsible for relevant losses in livestock, causing abortions, reducing milk production and leading to depreciation of activity, in addition to impair national and international trade of animals and their products (BERNUES; MANRIQUE; MAZA, 1997; MCDERMOTT; GRACE; ZINSSTAG, 2013). Brucellosis is also a significant public health concern (PAPPAS; AKRITIDIS; BOSILKOVSKI; TSIANOS, 2005), since, in humans, the disease often progress to chronicity, affecting a large number of systems, harming the quality of life of infected individuals (SINGH; KHATKAR; AULAKH; GILL *et al.*, 2018).

Due to the great damages in human and animal health caused by the disease, *Brucella* spp. has been studied over decades by several approaches involving microbiology, immunology, molecular biology, and genetics. The progress achieved in researches regarding this pathogen has been accelerated in recent years, due to the increase availability of whole genome sequences

in the main scientific online platforms (WHATMORE; FOSTER; EVOLUTION, 2021). The greatest amount of genomes from different isolates of the same pathogen has made it possible to carry out comparative genetic studies using a pan-genomic approach (SIVASHANKARI; SHANMUGHAVEL, 2007). Pan-genome analyzes provides a better understanding of phylogenomic evolution, enables the identification of genetic clusters common to all isolates of a given species, as well as clarifies the adaptation mechanism of pathogens in the host organism (COSTA; GUIMARÃES; SILVA; SOARES *et al.*, 2020; RUST; MONGIN; BIRNEY, 2002; SNIPEN; ALMØY; USSERY, 2009). Throughout the identification of the core-genome (genes present in all isolates), shared-genome (genes absent in one or more strains) and singletons (genes unique to each strain), the pan-genome analysis can be applied in the development of new vaccines and in the identification of resistance mechanisms hitherto unknown in some species (CARLOS GUIMARAES; BENEVIDES DE JESUS; VINICIUS CANARIO VIANA; SILVA *et al.*, 2015), providing advances in the treatment, control and prevention of pathogens, such as *B. abortus* (BARH; SOARES; TIWARI; AZEVEDO, 2020).

However, the quality of genomes included in the pan-genome analysis should be checked carefully, since the insertion of contaminations/incompleteness assembles may underestimate the core genes and compromise the identification of singletons, which generates bias on the results of the pan-genome and genetic repertory prediction for the assessed species (YANG; GAO, 2022). Indeed, it is well stated that reliable pan-genome results depends on high-quality genomes inputs and that incomplete assemblies or inconsistent gene annotations should be excluded from pan-genome analysis at the early stage of data processing, improving its efficiency, and ultimately contributing to a better understanding of genome function and evolution (LI; YIN, 2022; WU; WANG; GAO, 2021).

Until March 2023, 256 *B. abortus* genomes from strains isolated in different regions of the world were available at National Center for Biotechnology Information (NCBI). This study aimed to perform a pan-genome analysis of *B. abortus*, comparing the results of strains isolated from cattle in Brazil, alone or in combination with whole genomes available NCBI platform, assessing the impact of the genome quality in the results obtained.

## MATERIALS AND METHODS

### Brazilian strains

Fifty-three *B. abortus* strains previously isolated from naturally infected cattle in different Brazilian states, between 1977 and 2008 (MINHARRO; MOL; DORNELES; PAULETTI *et al.*, 2013), were used. All bacteriological (ALTON; JONES; ANGUS; VERGER, 1988) and molecular analyses (BAILY; KRAHN; DRASAR; STOKER, 1992; BRICKER; HALLING, 1994) to confirm the species and biovar of the strains (ALTON; JONES; ANGUS; VERGER, 1988) were previously performed (Table 1) (MINHARRO; MOL; DORNELES; PAULETTI *et al.*, 2013). Additionally, the strains were also classified into genotypes by means of MLVA16 analysis (LE FLÈCHE; JACQUES; GRAYON; AL DAHOUK *et al.*, 2006; MINHARRO; MOL; DORNELES; PAULETTI *et al.*, 2013).

The 53 Brazilian isolates were also previously tested for antimicrobial susceptibility against drugs commonly used in the treatment of human brucellosis: ciprofloxacin, doxycycline, streptomycin, gentamicin, ofloxacin, rifampicin and trimethoprim plus sulfamethoxazole (1 part trimethoprim to 19 parts of sulfamethoxazole) (BARBOSA PAULETTI; REINATO STYNEN; PINTO DA SILVA MOL; SELES DORNELES *et al.*, 2015). Data on isolation year, state, biovar, MLVA-16 genotype and antimicrobial susceptibility for each strain are shown in the Supplementary Table S1.

**Genome sequencing, assembly and curation**

The process of sequence, assembly and curation of the genomes from the 53 Brazilian strains were performed as previously described by PEREIRA; KATO; ARAÚJO; DA SILVA *et al.* (2023), resulting in 53 files with two closed chromosomes each. The genomes were deposited in the NCBI database, compiled under a single Bioproject: PRJNA750793. Data on epidemiological information, reference chosen for ordination, genome coverage and Biosample for each strain are provided at Supplementary Table S1.

**Download and quality assessment from NCBI genomes**

All available Refseq assemblies of *B. abortus* at NCBI were downloaded (09 March 2023) and when accessible, relevant metadata, such as date of collection, country of isolation and host species were retained. The downloaded genomes were evaluated by bbmap (BUSHNELL, 2014), and only those showing an overall size between 3.0 and 3.4 megabases, number of contigs < 250 and N50 > 125,000 were retained, according to parameters previously established for *Brucella* spp. (ORSINI; IANNI; ZINZULA, 2022). The assemblies were also analyzed imposing at least 98% of completeness and contamination lower than 10% using CheckM (PARKS; IMELFORT; SKENNERTON; HUGENHOLTZ *et al.*, 2015). The strains

with more than one genome deposited at NCBI had their duplicates removed by adopting the following criteria: smaller number of contigs and better genome integrity. Detailed information on the NCBI downloaded genomes and Brazilian genomes are shown in the Supplementary Table S2.

**Annotation and pan-genome characterization**

The assemblies approved in the adopted criteria were annotated using Prokka (SEEMANN, 2014) and the ".gff" files were used to predict the species pan-genome using Roary (default parameters) (PAGE; CUMMINS; HUNT; WONG *et al.*, 2015). The pan-genome analysis was performed using the same annotation files and the same parameters with two distinct groups: all the genomes available (Brazilian genomes + all NCBI downloaded genomes) or only high-quality genomes (Brazilian genomes + NCBI genomes that passed the established quality criteria).

The number of hypothetical proteins predicted by Prokka in each genome was obtained from the Roary output file. The core genome was defined as the genes present in 100% of the genomes included in the analysis. The accessory genome was divided into two groups: high frequency (HFA) - when present in 50% or more isolates - and low frequency (LFA), when present in less than 50% of the isolates (ORSINI; IANNI; ZINZULA, 2022). Singletons were considered genes unique to each genome, absent from the other components of the group (even if they occurred in paralogs, duplicated in the same strain).

**RESULTS**

**Genomes quality assessment**

The evaluation of the quality criteria resulted in the inclusion of 199 genomes, from the 256 assessed. Among the 57 excluded genomes, 38 did not have a size between 3.0 and 3.4 megabases, 3 had more than 250 contigs, 12 showed a N50 value greater than 125,000, 37 less than 98% of completeness, and 37 more than 10% of contamination. All Brazilian genomes passed in the adopted quality criteria. Also, 7 duplicates were excluded, 6 considering the first criterion (smaller amount of contigs) and 1 according to the second criterion (greater completeness). Among five parameters evaluated in the quality assessment, 36 (63.16%) genomes did not pass in three, 5 (8.77%) genomes in two, and 16 (28.07%) genomes (including the 7 duplicates) in one criterion. An overview of the quality assessment of all genomes are

shown in the Fig. 1 and detailed information on the five assessed parameter for each of the 309 genomes a in Supplementary file S3.

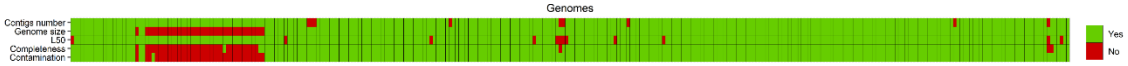


Figure 1: Heat map showing the results of the quality assessment according to the five used criteria for inclusion (yes) or exclusion (no) of the genome from the pan-genome analysis.

**Influence of low-quality B. abortus genomes on pan-genome analysis**

The adoption of a quality filter on the genomes led to a decrease of 41.30% of the pangenome (Table 1). Furthermore, it was observed that the number of genes in the core-genome tripled, while the number of singletons dropped by almost 2.5 times (Table 1). Nevertheless, the number of hypothetical proteins, as well as the number genes of the accessory genome (HFA and LFA), despite having showed a reduction in their absolute number, remained with very similar relative frequencies in relation to the total gene repertoire of the group (Table 1). The difference between the proportion of genes common to all isolates considering the two subgroups, with and without the quality criteria screening, can be seen in Fig. 2.

Table 1: Pan-genome analysis for the two subgroups of *Brucella abortus* genomes analyzed, with and without the adoption of quality criteria.

| Parameters | No filter | Quality selection |
|---|---|---|
| Number of genomes | 309 (256 + 53) | 252 (199 + 53) |
| Core-genome | 1273 (15.1%) | 2246 (45.3%) |
| Accessory genome (HFA ≥ 50%) | 1861 (22.0%) | 903 (18.2%) |
| Accessory genome (LFA < 50%) | 1684 (19.9%) | 915 (18.5%) |
| Singletons | 3624 (42.9%) | 891 (18.0%) |
| Pan-genome | 8442 | 4955 |
| Hypotetical proteins | 3840 (45.5%) | 2087 (42.1%) |

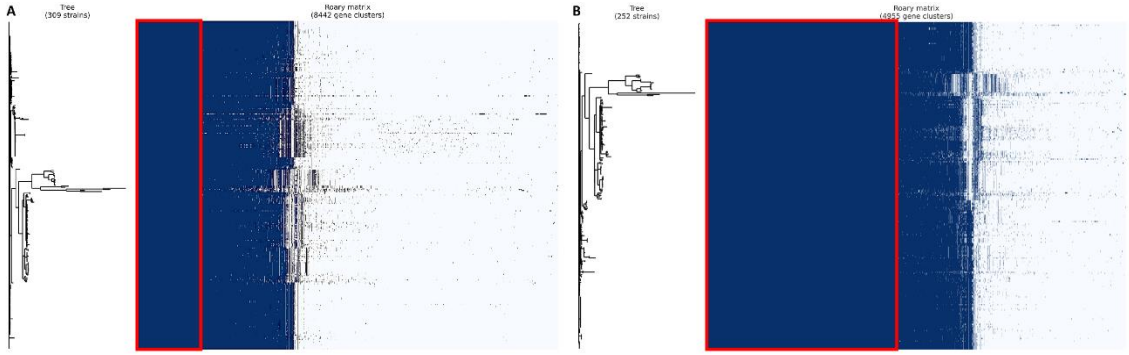


Figure 2: Phylogenetic tree and Roary gene presence and absence matrix for 309 *Brucella abortus* genomes not submitted to any quality assessment (A) and the same analysis for 252 *B. abortus* genomes selected after a quality analysis (B). The highlighted groups (red) represent the core-genes.

**Comparison between Brazil and NCBI pan-genome subgroups**

The Brazilian genomes represented 21.0% of the strains included in the *B. abortus* pan-genome analysis. However, the Brazilian isolates added only 34 genes in the total pan-genome, which correspond to 0.7% of the gene repertoire of the species.

However, an inversely proportional behavior was observed in the distribution of genes in the core-genome and in the singletons between the Brazil and NCBI subgroups. The first had almost twice as many genes common to all its components and approximately seven times fewer genes exclusive to only one genome. In fact, this difference can also be observed at the accessory gene from the Brazil subgroup compared to "NCBI" and "Total" groups, since the number of accessory genes present in 50% or more of the isolates (HFA) was twice of the present in less than 50% of the isolates (LFA). This behavior of the accessory genome was not observed in the "NCBI" and "Total" groups, in which the HFA and LFA genes exhibited a homogeneous distribution among themselves (Table 2).

Table 2: Pan-genome parameters for the Brazilian, NCBI and all *Brucella abortus* genomes after the quality selection.

| Parameters | Brazil | NCBI | Total (Brazil + NCBI) |
|---|---|---|---|
| Number of genomes | 53 (21.0%) | 199 (79.0%) | 252 (100%) |
| Core-genome | 2,934 (87.3%) | 2,289 (46.5%) | 2,246 (45.3%) |
| Accessory genome (HFA ≥ 50%) | 221 (6.6%) | 854 (17.4%) | 903 (18.2%) |
| Accessory genome (LFA < 50%) | 123 (3.7%) | 902 (18.3%) | 915 (18.5%) |
| Singletons | 83 (2.5%) | 876 (17.8%) | 891 (18.0%) |

| | | | |
|---|---|---|---|
| Pan-genome | 3,361 | 4,921 | 4,955 |
| Hypotetical proteins | 1,252 (37.3%) | 2,074 (42.1%) | 2,087 (42.1%) |

**DISCUSSION**

The advances in sequencing technologies and the development of bioinformatics tools and approaches have generated an exponential growth of bacterial genomic data, making it possible to perform pan-genome analysis to better understand and characterize the genome of several species (RUST; MONGIN; BIRNEY, 2002; SNIPEN; ALMØY; USSERY, 2009). The comparative analyzes allowed an estimation of the impact in quantitative levels of the addition of low quality assembles in the genomic characterization of *B. abortus*. Furthermore, it was possible to identify a large number of hypothetical proteins present in the pan-genome, demonstrating how much progress is still need in the characterization of the gene repertoire and encoded proteins of this species. Finally, it was possible to detect an increment of 34 genes caused by the addition of the 53 Brazilian strains in the pan-genome of the species.

The adoption of minimum quality parameters was fundamental in the analysis for adequate characterization of the species pan-genome, mainly of the core-genome. In a preliminary analysis, without any filter, it was identified that the pan-genome corresponded to only 15% of the total number of genes, which is a bias, since the *B. abortus* genome is extremely conserved and clonal, with practically no lateral acquisition of genes (SUÁREZ-ESQUIVEL; CHAVES-OLARTE; MORENO; GUZMÁN-VERRI, 2020). When the quality filter was applied, the proportion of genes shared by all species increased by more than three times. This can be explained by the fact that the inclusion of a single segmented or incomplete genome can lead to the incorrect conclusion that that strain does not have a certain gene, when, in fact, it was just not detected due to poor assembly quality. This incorrectly exclude this gene from the group called core-genome, causing a misinterpretation regarding the conservation of this pathogen's DNA (WU; WANG; GAO, 2021).

Similarly, the inclusion of assemblies of contaminated sequencing directly affects the number of singletons and the reliability of the analyses (YANG; GAO, 2022). While the fragmentation and incompleteness of genomes decreases the number of genes common to all the analyzed strains, the insertion of contaminated genomes increases the number of unique genes, as it leads to a misinterpretation that some *B. abortus* isolates contain genes that do not actually belong to the species. Indeed, it has been discussed in the literature that no matter how small is the percentages of contamination in the genomes deposited on the NCBI platform

(estimated at 1.25% of bacterial genomes Refseq) (LUPO; VAN VLIERBERGHE; VANDERSCHUREN; KERFF *et al.*, 2021), a single genome of low quality when inserted in pan-genome population analyzes is capable of generating false interpretations in all results and will ultimately result in an overestimation of the pan-genome of the species (LI; YIN, 2022), as observed in the present study.

Our results also enabled the identification of a high proportion of hypothetical proteins, representing more than 40% of the coding sequences of the species. Although the number of *Brucella* spp. genomes sequenced and publicly available on platforms has grown exponentially (WHATMORE; FOSTER; EVOLUTION, 2021), it is necessary to emphasize that not all questions can be answered by this tool. Despite the prediction algorithms, many proteins in *B. abortus* have their functions still unknown, which represents a great challenge for advances not only in the areas, such as antimicrobial resistance, reverse vaccinology and functional genomics, but also to general biology (GALPERIN, 2001). In this sense, it is worth to note that filtering the studied strains by quality was responsible for a decrease of only 3% in the number of hypothetical proteins, remaining 42% of genes with unknown function that need to be further explored and characterized.

The addition of the Brazilian genomes to the pan-genome caused a small increase in the core-genome, evidencing that the genomes circulating in the country are extremely clonal and conserved, as can be confirmed by the high percentage of genes common to all Brazilian isolates. This finding could be supported by the epidemiological context of the genomes included in each subset: while the 53 Brazilian genomes were isolated from a single country, between 1977 and 2008, the 199 genomes available in the NCBI were isolated from at least three continents, including countries as China, India, Italy, South Korea and United States, from 1919 to 2019 (data available from PATRIC platform – https://patricbrc.org/). This same restricted epidemiological context regarding Brazilian genomes may explain the low percentage of singletons in the "pan-Brazil" subgroup. It is worth mentioning that all Brazilian genomes were isolated from cattle, which have little or no mobility with other countries (decreasing the genetic diversity of circulating isolates), different from other the hosts, as humans.

The main goal of a pan-genome is understanding the strain diversity of a species through comparisons at the genomic level. The inclusion of the 53 Brazilian genomes in the analysis was responsible for adding 34 genes to the "Total" pan-genome (4,955 genes), suggesting that the Brazilian strains have a genetic repertoire, which until then had not been observed in *B. abortus* from other regions of the world. These advances in the knowledge about the genetic

composition of a given bacterial species, as *B. abortus*, will certainly contribute to studies involving metabolic pathways, evolutionary history and molecular fingerprint targets for epidemiological studies, among others (ROULI; MERHEJ; FOURNIER; RAOULT, 2015). Studies of functional groups of genes, phylogenomic, identification of genomic islands and implementation of machine learning methods to elucidate genetic determinants of antimicrobial resistance are only some of the possible approaches that can be conducted from this repertoire of data generated by the pan-genome. Furthermore, these analyzes have also a great potential to increase the current knowledge about *B. abortus*, especially when integrated with the phenotypic and epidemiological data already available on the isolates.

In conclusion, the quality curation of assemblies before the performance of a pan-genome analysis proved to be an essential step for an accurate characterization of the genetic repertoire of *B. abortus*. Many genes of this species have not yet been characterized, as can be seen by the high number of hypothetical proteins in the pan-genome, even after applying the quality parameters. This is a field that still needs to be explored in *B. abortus* for a better understanding of the genetic mechanisms that regulate the biological processes of this organism. The inclusion of the 53 complete genomes from strains isolated in Brazil was responsible for adding 34 genes to the known genetic repertoire of *B. abortus*.

## SUPPLEMENTARY FILES

https://drive.google.com/drive/folders/1vOtahFFRv7pk0kWVbLRhQkvFIDeOEzKy?usp=drive_link

S1 - Data on isolation year, state, biovar, MLVA-16 genotype and antimicrobial susceptibility for each Brazilian strain.

S2 - Detailed information of NCBI downloaded genomes and Brazilian genomes.

S3 - Five assessed parameter for each of the 309 genomes assessed in this study.

## ACKNOWLEDGMENTS

This study was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq [423557/2016-5], Fundação de Amparo à Pesquisa do Estado de Minas Gerais – Fapemig [APQ-00283/2016 and RED-00132-22], Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes and Network of Omics Sciences (RECOM). DO'C

received a travel grant from CAPES/PrInt/UFLA [52/2019]. CRP received a scholarship from CAPES/PrInt/UFLA [01/2020]. EMSD is also grateful to CNPq for her fellowship.

## REFERENCES

ABOUT, F.; PASTRE, T.; BOUTROU, M.; MARTINEZ, A. Y. *et al.* Novel Species of *Brucella* Causing Human Brucellosis, French Guiana. **Emerg Infect Dis**, 29, n. 2, p. 333-340, Feb 2023.

ALTON, G. G.; JONES, L. M.; ANGUS, R.; VERGER, J. **Techniques for the brucellosis laboratory**. Institut National de la recherche Agronomique (INRA), 1988. 2738000428.

BAILY, G.; KRAHN, J.; DRASAR, B.; STOKER, N. Detection of *Brucella melitensis* and *Brucella abortus* by DNA amplification. **Tropical Medicine & International Health**, 95, n. 4, p. 271-275, 1992.

BARBOSA PAULETTI, R.; REINATO STYNEN, A. P.; PINTO DA SILVA MOL, J.; SELES DORNELES, E. M. *et al.* Reduced Susceptibility to Rifampicin and Resistance to Multiple Antimicrobial Agents among *Brucella abortus* Isolates from Cattle in Brazil. **PLOS ONE**, 10, n. 7, p. e0132532, 2015.

BARH, D.; SOARES, S.; TIWARI, S.; AZEVEDO, V. Pan-genomics: Applications, Challenges, and Future Prospects. *In*: : **Academic Press**, 2020. p. iv.

BERNUES, A.; MANRIQUE, E.; MAZA, M. Economic evaluation of bovine brucellosis and tuberculosis eradication programmes in a mountain area of Spain. **Preventive veterinary medicine**, 30, n. 2, p. 137-149, 1997.

BRICKER, B. J.; HALLING, S. M. Differentiation of *Brucella abortus* bv. 1, 2, and 4, *Brucella melitensis*, *Brucella ovis*, and *Brucella suis* bv. 1 by PCR. **Journal of clinical microbiology**, 32, n. 11, p. 2660-2666, 1994.

BUSHNELL, B. **BBMap: a fast, accurate, splice-aware aligner**. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States). 2014.

CARLOS GUIMARAES, L.; BENEVIDES DE JESUS, L.; VINICIUS CANARIO VIANA, M.; SILVA, A. *et al.* Inside the pan-genome-methods and software overview. **Current genomics**, 16, n. 4, p. 245-252, 2015.

COSTA, S. S.; GUIMARÃES, L. C.; SILVA, A.; SOARES, S. C. *et al.* First Steps in the Analysis of Prokaryotic Pan-Genomes. **Bioinform Biol Insights**, 14, p. 1177932220938064, 2020.

FRANC, K.; KRECEK, R.; HÄSLER, B.; ARENAS-GAMBOA, A. Brucellosis remains a neglected disease in the developing world: a call for interdisciplinary action. **BMC public health**, 18, n. 1, p. 1-9, 2018.

GALPERIN, M. Y. Conserved 'hypothetical' proteins: new hints and new puzzles. **Comp Funct Genomics**, 2, n. 1, p. 14-18, 2001.

KIIZA, D.; DENAGAMAGE, T.; SERRA, R.; MAUNSELL, F. *et al.* A systematic review of economic assessments for brucellosis control interventions in livestock populations. **Prev Vet Med**, 213, p. 105878, Apr 2023.

LE FLÈCHE, P.; JACQUES, I.; GRAYON, M.; AL DAHOUK, S. *et al.* Evaluation and selection of tandem repeat loci for a *Brucella* MLVA typing assay. **BMC microbiology**, 6, n. 1, p. 9, 2006.

LI, T.; YIN, Y. Critical assessment of pan-genomic analysis of metagenome-assembled genomes. **Briefings in Bioinformatics**, 23, n. 6, 2022.

LUPO, V.; VAN VLIERBERGHE, M.; VANDERSCHUREN, H.; KERFF, F. *et al.* Contamination in Reference Sequence Databases: Time for Divide-and-Rule Tactics. **Front Microbiol**, 12, p. 755101, 2021.

MCDERMOTT, J.; GRACE, D.; ZINSSTAG, J. Economics of brucellosis impact and control in low-income countries. **Rev Sci Tech**, 32, n. 1, p. 249-261, 2013.

MINHARRO, S.; MOL, J. P. S.; DORNELES, E. M.; PAULETTI, R. B. *et al.* Biotyping and genotyping (MLVA16) of *Brucella abortus* isolated from cattle in Brazil, 1977 to 2008. **PLoS One**, 8, n. 12, p. e81152, 2013.

ORSINI, M.; IANNI, A.; ZINZULA, L. *Brucella ceti* and *Brucella pinnipedialis* genome characterization unveils genetic features that highlight their zoonotic potential. 11, n. 5, p. e1329, 2022.

PAGE, A. J.; CUMMINS, C. A.; HUNT, M.; WONG, V. K. *et al.* Roary: rapid large-scale prokaryote pan genome analysis. 31, n. 22, p. 3691-3693, 2015.

PAPPAS, G.; AKRITIDIS, N.; BOSILKOVSKI, M.; TSIANOS, E. Brucellosis. **New England Journal of Medicine**, 352, n. 22, p. 2325-2336, 2005.

PAPPAS, G.; PAPADIMITRIOU, P.; AKRITIDIS, N.; CHRISTOU, L. *et al.* The new global map of human brucellosis. **The Lancet Infectious Diseases**, 6, n. 2, p. 91-99, 2006.

PARKS, D. H.; IMELFORT, M.; SKENNERTON, C. T.; HUGENHOLTZ, P. *et al.* CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. 25, n. 7, p. 1043-1055, 2015.

PEREIRA, C. R.; KATO, R. B.; ARAÚJO, F. A.; DA SILVA, A. L. *et al.* Genomic investigation of antimicrobial resistance in *Brucella abortus* strains isolated from cattle in Brazil. **Gene Reports**, 31, p. 101777, 2023/06/01/ 2023.

ROULI, L.; MERHEJ, V.; FOURNIER, P.-E.; RAOULT, D. The bacterial pangenome as a new tool for analysing pathogenic bacteria. **New microbes and new infections**, 7, p. 72-85, 2015.

RUST, A. G.; MONGIN, E.; BIRNEY, E. Genome annotation techniques: new approaches and challenges. **Drug discovery today**, 7, n. 11, p. S70-S76, 2002.

SEEMANN, T. J. B. Prokka: rapid prokaryotic genome annotation. 30, n. 14, p. 2068-2069, 2014.

SINGH, B. B.; KHATKAR, M. S.; AULAKH, R. S.; GILL, J. P. S. *et al.* Estimation of the health and economic burden of human brucellosis in India. **Preventive Veterinary Medicine**, 154, p. 148-155, 2018/06/01/ 2018.

SIVASHANKARI, S.; SHANMUGHAVEL, P. Comparative genomics-a perspective. **Bioinformation**, 1, n. 9, p. 376, 2007.

SNIPEN, L.; ALMØY, T.; USSERY, D. W. Microbial comparative pan-genomics using binomial mixture models. **BMC genomics**, 10, n. 1, p. 1-8, 2009.

SUÁREZ-ESQUIVEL, M.; CHAVES-OLARTE, E.; MORENO, E.; GUZMÁN-VERRI, C. *Brucella* Genomics: Macro and Micro Evolution. 21, n. 20, p. 7749, 2020.

WHATMORE, A. M.; FOSTER, J. T. J. I., Genetics; EVOLUTION. Emerging diversity and ongoing expansion of the genus *Brucella*. 92, p. 104865, 2021.

WU, H.; WANG, D.; GAO, F. Toward a high-quality pan-genome landscape of *Bacillus subtilis* by removal of confounding strains. **Brief Bioinform**, 22, n. 2, p. 1951-1971, Mar 22 2021.

YANG, T.; GAO, F. High-quality pan-genome of *Escherichia coli* generated by excluding confounding and highly similar strains reveals an association between unique gene clusters and genomic islands. **Briefings in Bioinformatics**, 23, n. 4, 2022.

# CHAPTER 5 - First report and whole-genome sequencing of *Pseudochrobactrum saccharolyticum* in Latin America

Carine Rodrigues Pereira[1], Thiago de Jesus Sousa[2], Alessandra Lima da Silva[2], Roselane Gonçalves dos Santos[2], Sílvia Minharro[3], Dirceia Aparecida Costa Custódio[1], Derek J. Pickard[4], David O'Callaghan[56], Jeffrey T. Foster[7], Siomar de Castro Soares[8], Rommel Thiago Juca Ramos[9], Aristóteles Góes-Neto[2], Mateus Matiuzzi da Costa[10], Andrey Pereira Lage[11], Vasco Azevedo[2], Elaine Maria Seles Dorneles[1*]

[1]Departamento de Medicina Veterinária, Faculdade de Zootecnia e Medicina Veterinária, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil

[2]Departamento de Genética, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

[3]Centro de Ciência da Saúde - Medicina - Araguaína, Universidade Federal de Tocantins, Tocantins, Brazil

[4]Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, University of Cambridge, Cambridge, United Kingdom

[5]IVBIC, INSERM, Universite de Montpellier, Nimes, France

[6]CNR Brucella, Laboratoire de Microbiologie, CHU Nimes, Nimes, France.

[7]Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, Arizona, USA

[8]Departamento de Microbiologia, Imunologia e Parasitologia, Instituto de Ciências Biológicas e Ciências Naturais, Universidade Federal do Triângulo Mineiro, Uberaba, Minas Gerais, Brazil

[9]Instituto de Ciências Biológicas, Centro de Genômica e Biologia de Sistemas, Universidade Federal do Pará, Belém, Pará, Brazil

[10]Universidade Federal do Vale do São Francisco, Departamento de Zootecnia, Petrolina, Pernambuco, Brazil

[11]Departamento de Medicina Veterinária Preventiva, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

*Corresponding author: +55 (35) 3829-1712 – elaine.dorneles@ufla.br – Universidade Federal de Lavras Caixa Postal 3037 – CEP: 37200-900 – Lavras – MG – Brazil

**ABSTRACT:**

The *Brucellaceae* family comprises microorganisms similar both phenotypically and genotypically, making it difficult to identify the etiological agent of these infections. This study reports the first isolation, identification, and characterization of *Pseudochrobactrum saccharolyticum* (strain 115) from Latin America. Strain 115 was isolated in 2007 from a bovine in Brazil and was initially classified as *Brucella* spp. by classical microbiological tests and *bcsp31* PCR. The antimicrobial susceptibility of strain 115 was tested against drugs used to treat human brucellosis by minimal inhibitory concentration test. Subsequently, the whole genome of the strain was sequenced, assembled, and characterized. Phylogenetic trees built from *16S rRNA* and *recA* gene sequences enabled the classification of strain 115 as *Pseudochrobactrum* spp. Phylogenomic analysis using Single Nucleotide Polymorphisms and Average Nucleotide Identity allowed the classification of the strain as *P. saccharolyticum*. Additionally, a Tetra Correlation Search identified one related genome from the same species, which was compared with strain 115 by analyzing genomic islands. This is the first identification and whole-genome sequence of *P. saccharolyticum* in Latin America and highlights a challenge in the diagnosis of bovine brucellosis, which could be solved by including the sequencing of *16S* and *recA* genes in routine diagnostics.

**KEYWORDS:** *Brucellaceae*; phylogenetic; antimicrobial resistance

## INTRODUCTION

The genera *Brucella* and *Pseudochrobactrum* are part of the *Brucellaceae* family (ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.*, 2020). Until 2020, the genus *Brucella* comprised 12 species existing largely in association with an animal host in both wild and domesticated animals (WHATMORE; FOSTER, 2021). The genus is disseminated throughout the world and several species are responsible for significant losses in animal production and pose substantial public health burdens (SCHOLZ; BANAI; CLOECKAERT; KÄMPFER *et al.*, 2015). Recently, the International Committee on Systematics of Prokaryotes reclassified the genus *Ochrobactrum* into *Brucella*, (HÖRDT; LÓPEZ; MEIER-KOLTHOFF; SCHLEUNING *et al.*, 2020) as well as standardized the new names of each of the 18 known *Brucella* species and their type strains (OREN; GARRITY, 2020). Different from the "traditional" *Brucella* species, *Ochrobactrum* spp. are often isolated from the environment and considered as an opportunistic pathogen, affecting animals and humans (ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.*, 2020; VILA; PAGELLA; BELLO; VICENTE,

2016). These medical and epidemiological differences between *Brucella* spp. and *Ochrobactrum* spp. together with important differences in genomic traits and pathogenicity have been used as arguments to maintain these two genera apart (MORENO; BLASCO; LETESSON; GORVEL *et al.*, 2022).

One genus most recently described in the family *Brucellaceae* is *Pseudochrobactrum*, which includes five species: *P. saccharolyticum* (first isolated from an industrial glue), *P. asaccharolyticum* (isolated from a knee aspirate of a 66-year-old man) (KÄMPFER; ROSSELLÓ-MORA; SCHOLZ; WELINDER-OLSSON *et al.*, 2006), *P. kiredjianiae* (isolated from stainless-steel vent covers in a seafood processing plant) (KÄMPFER; SCHOLZ; HUBER; THUMMES *et al.*, 2007), *P. lubricantis* (isolated from a metal-working fluid in a metal-processing company) (KÄMPFER; HUBER; LODDERS; WARFOLOMEOW *et al.*, 2009), and *P. algeriensis* (isolated from cattle lymph nodes) (LOPERENA-BARBER; KHAMES; LECLERCQ; ZYGMUNT *et al.*, 2022).

The genera that comprise the *Brucellaceae* family are genetically very closely related (ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.*, 2020). Besides this close genetic relationship, the similar clinical presentation of infections and frequent cross-reactions in diagnostic tests between these two genera often leads to issues in the proper identification of these microorganisms, from both animals and humans. Indeed, the misdiagnosis of pathogens of *Brucellaceae* has already been reported in which 13 isolates of veterinary origin submitted to a laboratory as suspect *Brucella* spp. were later confirmed as *Pseudochrobactrum* spp., indicating these related bacteria are also present in ruminant samples (ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.*, 2020; KÄMPFER; ROSSELLÓ-MORA; SCHOLZ; WELINDER-OLSSON *et al.*, 2006). Usually, whole-genome sequencing followed by phylogenomic analyses are used to distinguish between those two *Brucellaceae* genera and are particularly useful for the correct classification of isolates that do not present typical patterns in microbiological and molecular tests.

The present study reports the first isolation and identification of *P. saccharolyticum* in Latin America (strain 115), describes its phenotypic and genotypic features, and classifies this microorganism at phylogenetic and phylogenomic levels.

## MATERIAL AND METHODS

### Strain isolation, culture, and identification

Strain 115 was isolated from specimens collected at the slaughter of a cow in February 2007, in Ituiutaba, Minas Gerais, Brazil. Mammary lymph node and material from a tarsus lesion were pooled, macerated, and inoculated into tryptose broth supplemented with Farrell's antimicrobial mixture. After seven days incubation at 37 ºC in 5% $CO_2$ , 100 µL of broth were inoculated onto tryptose agar plates supplemented with Farrell's antimicrobial mixture which were incubated at the same conditions for 48h (ALTON; JONES; ANGUS; VERGER, 1988). Because the suspected diagnosis was bovine brucellosis, phenotypic identification was performed using standard methods (ALTON; JONES; ANGUS; VERGER, 1988).

The isolate was cultivated on tryptose agar plates for 48 hours in 5% $CO_2$ at 37 ºC, and the colonies were resuspended in 100 µL TE buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA) and inactivated at 85 ºC for 2 hours. According to the manufacturer's specification, the DNA was extracted using the Wizard Genomic DNA Purification Kit. *Brucella* genus-specific PCR (BAILY; KRAHN; DRASAR; STOKER, 1992) and AMOS-PCR (BRICKER; HALLING, 1994) were also used to try to identify the strain. Reference strains *B. abortus* biovar 4 292 = ATCC 23451, *B. abortus* biovar 1 544 = ATCC 23448$^T$, *B. abortus* biovar 1 2308, *B. abortus* biovar 1 S19, *B. abortus* biovar 1 RB51, *B. melitensis* biovar 1 16M = ATCC 23456T, *B. ovis* Reo 198, and *B. suis* 1330 (ATCC 23444) were used as controls in different microbiological and molecular tests.

Strain 115 was also analyzed in matrix-assisted laser desorption ionization (MALDI)-time of flight (TOF) mass spectrometry (MS) using a MALDI Biotyper. For this purpose, the isolates were grown on BHI (Brain Heart Infusion) agar for 24 hours at 37 °C and processed as previously described (ASSIS; PEREIRA; ZEGARRA; TAVARES *et al.*, 2017).

**Antimicrobial susceptibility tests**

Minimum Inhibitory Concentration (MIC) were determined using the agar dilution method according to the Clinical and Laboratory Standards Institute (CLSI), as previously described (BARBOSA PAULETTI; STYNEN; MOL; DORNELES *et al.*, 2015).

**Genome sequencing, assembling, annotation, and characterization**

Whole-genome sequencing was performed on the Illumina HiSeq 2500 platform with 2 x 151 bp paired-end reads. The quality of DNA sequences was assessed using FastQC raw data (ANDREWS; KRUEGER; SECONDS-PICHON; BIGGINS *et al.*, 2010), and the genome was assembled using Newbler v. 2.9 (MARGULIES; EGHOLM; ALTMAN; ATTIYA *et al.*, 2005). The scaffolds resulting from the assembly were ordered in MeDuSa (BOSI; DONATI;

GALARDINI; BRUNETTI *et al.*, 2015), using *P. saccharolyticum* CCUG 33852$^T$ as the reference, which was identified in Tetra Correlation Search (TCS) and analyzed in Average Nucleotide Identity based on BLAST (ANIb), with more than 95% of nucleotides aligned in JSpecies platform (http://jspecies.ribohost.com/jspeciesws/). After ordering, the scaffolds gaps were closed in GFinisher (GUIZELINI; RAITTZ; CRUZ; SOUZA *et al.*, 2016). Curation of genomes coverage and quality control was carried out using CLC Genomic Workbench software (Qiagen, USA) and annotation was performed using PROKKA (SEEMANN, 2014).

The draft genome of strain 115 was submitted to the Comprehensive Genome Analysis (CGA) on the PATRIC platform (https://www.patricbrc.org/) and analyzed in GIPSy (SOARES; GEYIK; RAMOS; DE SÁ *et al.*, 2016) to predict genomic islands, using *B. abortus* 2308 as a reference. It was compared with *P. saccharolyticum* CCUG 33852$^T$ using BLAST Ring Image Generator (BRIG) (ALIKHAN; PETTY; ZAKOUR; BEATSON, 2011). The synteny of the assemblies of these three strains were also compared with Mauve (DARLING; MAU; BLATTNER; PERNA, 2004).

**Phylogenetic analysis**

The *16S rRNA* and *recA* genes of strain 115 were extracted after PROKKA annotation. The genes were separately aligned using Muscle (Multiple Sequence Comparison by Log-Expectation) with 22 strains of three different bacterial genera obtained from KÄMPFER; HUBER; LODDERS; WARFOLOMEOW *et al.* (2009) (Table 1), in MEGAX v.11 (KUMAR; STECHER; LI; KNYAZ *et al.*, 2018). The updated classification of genus, species and type strains followed OREN e GARRITY (2020). A phylogenetic tree, rooted with *Bartonella* spp., was constructed using a neighbor-joining reconstruction method with the Kimura two-parameter correction and 1000 bootstrap iterations, as previously described (KÄMPFER; HUBER; LODDERS; WARFOLOMEOW *et al.*, 2009; SCHOLZ; TOMASO; DAHOUK; WITTE *et al.*, 2006).

Table 1: Strains and National Center for Biotechnology Information (NCBI) accession numbers of representatives from the *Brucellaceae* family and outgroup *Bartonella* spp. strains used in phylogenetic analysis of *16S* rRNA and *recA* genes.

| **Strain** | **NCBI accession numbers** | |
|---|---|---|
| | ***16S*** | ***recA*** |
| *Bartonella henselae* Houston-1$^T$ | - | NC_005956 |
| *Bartonella henselae* M40SHD | DQ645426 | - |
| *Brucella abortus* NCTC 10093$^T$ | AM158979 | AM113730 |
| *Brucella melitensis* NCTC 10094$^T$ | - | AM087912 |

| | | |
|---|---|---|
| *Brucella rhizosphaerae* PR 17$^T$ | NR_042600 | AM422876 |
| *Brucella anthropi* CCUG 38531 | AM114405 | AM087904 |
| *Brucella anthropi* CIP 82.115 $^T$ | AM114398 | AM087899 |
| *Brucella gallinifaecis* Iso 196$^T$ | AJ519939 | AM422962 |
| *Brucella grignonensis* OgA9a$^T$ | AM490619 | AM422960 |
| *Brucella intermedia* CNS 2-75$^T$ | AM114411 | AM087913 |
| *Brucella intermedia* CCUG 43465 | AM490610 | AM422879 |
| *Brucella lupini* LMG 20667$^T$ | - | AM113737 |
| *Brucella oryzae* DSM 17471$^T$ | AM041247 | AM263421 |
| *Brucella pseudintermedia* ADV31$^T$ | DQ365921 | KF866348 |
| *Brucella pseudogrignonensis* CCUG 30717$^T$ | NR_042589 | AM422877 |
| *Brucella thiophenivorans* CCM 7492$^T$ | AM490617 | AM422872 |
| *Brucella tritici* SCII24$^T$ | AM114402 | AM087914 |
| *Pseudochrobactrum algeriensis* C130915_07$^T$ | MZ227818 | GCF_018436245.1[a] |
| *Pseudochrobactrum asaccharolyticum* CCUG 46016$^T$ | AM180485 | AM118081 |
| *Pseudochrobactrum kiredjianiae* CCUG 49584$^T$ | AM263420 | AM263419 |
| *Pseudochrobactrum lubricantis* KSS 7.8$^T$ | FM209496 | FM211811 |
| *Pseudochrobactrum saccharolyticum* CCUG 33852$^T$ | AM180484 | AM118082 |
| *Pseudochrobactrum saccharolyticum* strain 115 | This study[b] | This study[b] |

[a]Extracted from the whole-genome sequence; [b]Extracted from the whole-genome sequence (SAMN17006014)

**Phylogenomic analysis**

The sequence read archives and genome assemblies of the *Pseudochrobactrum* spp. type strains were obtained from ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.* (2020) and LOPERENA-BARBER; KHAMES; LECLERCQ; ZYGMUNT *et al.* (2022) (Table 2). These sequences were used to build a phylogenomic tree, rooted with *B. abortus* 2308, using the Northern Arizona SNP Pipeline (NASP) (SAHL; LEMMER; TRAVIS; SCHUPP *et al.*, 2016) and IQ-TREE (NGUYEN; SCHMIDT; VON HAESELER; MINH *et al.*, 2015) with default parameters. The phylogenomic tree was customized in iTOL (LETUNIC; BORK, 2007). An ANI analysis was performed using FastANI (JAIN; RODRIGUEZ-R; PHILLIPPY; KONSTANTINIDIS *et al.*, 2018) after assembling the reads provided by ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.* (2020) with SPAdes (BANKEVICH; NURK; ANTIPOV; GUREVICH *et al.*, 2012). The heatmap was plotted in R, using the packages reshape2 (WICKHAM, 2007), ComplexHeatmap (GU; EILS; SCHLESNER, 2016) and gplots (WARNES; BOLKER; BONEBAKKER; GENTLEMAN *et al.*, 2016) (https://github.com/spencer411/FastANI_heatmap).

Table 2: Strains and National Center for Biotechnology Information (NCBI) accession numbers of representatives from the *Pseudochrobactrum* genus and outgroup *Brucella abortus* strain used in phylogenomic and ANI analysis.

| Strains | NCBI accession numbers |
|---|---|
| *Brucella abortus* 2308 | GCF_000054005.1 |
| *Pseudochrobactrum algeriensis* C130915_07$^T$ | NZ_CP075361.1 |
| *Pseudochrobactrum asaccharolyticum* CCUG 46016$^T$ | SRR12012451 |
| *Pseudochrobactrum kiredjianiae* CCUG 49584$^T$ | SRR12012449 |
| *Pseudochrobactrum lubricantis* KSS 7.8$^T$ | SRR12012448 |
| *Pseudochrobactrum saccharolyticum* CCUG 33852$^T$ | SRR12012447 |
| *Pseudochrobactrum saccharolyticum* strain 115 | This study[a] |

[a] whole-genome sequence (SAMN17006014) and sequence read archive (SRR19103617)

## RESULTS

Strain 115 was characterized as Gram-negative and exhibited the following results in phenotypical tests: nitrate reduction (+), catalase (+), oxidase (+), agglutination in acriflavine (+), urease (+), production of hydrogen sulfide ($H_2S$) (+), citrate (-), the requirement for supplementary carbon dioxide ($CO_2$) (-), and motility (-). All biochemical and phenotypical tests and the genus-specific *Brucella* PCR were consistent with the classification of strain 115 as belonging to *Brucella* spp, however by AMOS-PCR no amplification was detected. The MIC value for each tested antimicrobial was as follows: ciprofloxacin (1 μg/mL), doxycycline (0.25 μg/mL), gentamicin (4 μg/mL), ofloxacin (1 μg/mL), rifampicin (8 μg/mL), streptomycin (>128 μg/mL) and sulfamethoxazole + trimethoprim (64 μg/mL). In the later stages of the study, strain 115 was identified in MALDI-TOF analysis as *Pseudochrobactrum asaccharolyticum*, with a score of 2,304 (Appendix A).

The scaffolds of strain 115 were submitted to the JSpecies platform and showed high similarity with *P. saccharolyticum* CCUG 33852$^T$ (ANIb – 98.02% and nucleotides aligned – 95.16%), which was used as the reference to order strain 115 in MeDuSa. After gap closure, the final draft contained 16 contigs, 4 persistent gaps, and a genome size of 3,876,228 bp. The reads and genome were deposited as a draft to the National Center for Biotechnology Information (NCBI) under SRA SRR19103617, BioProject PRJNA682717 and BioSample SAMN17006014, with 214x coverage and 98.75% of the reads mapped. The GC content was 52.55%. A total of 3,548 coding sequences (CDS) were predicted, after annotation. Two rRNA clusters of *5S*, *16S*, and *23S* were predicted, and 57 tRNA coding genes were detected.

A circular graphical display of the *P. saccharolyticum* strain 115 genome annotation and subsystems in CGA identified the location of eleven subsystems along the genome, as well as the position of CDS annotated with similarity to known antimicrobial resistance genes and virulence factors (Fig. 1 and Appendix B). Furthermore, a comparison between *P.*

*saccharolyticum* type strain CCUG 33852$^T$ and strain 115 in BRIG allowed the identification of some regions present/absent in strain CCUG 33852$^T$ and genomic islands were visualized in GIPSy. The genomic island predictions were classified as “normal” or “strong”, with functions related to metabolism (MI = 1), pathogenicity (PAI = 3), antimicrobial resistance (RI = 7), and miscellaneous islands (MSI = 5) (Fig. 2) (Appendix C). In addition, BRIG analysis identified inversions in the *P. saccharolyticum* 115 genome compared with the *P. saccharolyticum* type strain CCUG 33852$^T$, which was also visualized in Mauve (Appendix D). Phylogenetic analysis of the *16S* and *recA* genes allowed the identification of strain 115 at the genus level, being closest to the species *P. saccharolyticum* and *P. lubricantis* (Fig. 3). On the other hand, SNP analyses from whole genomes and ANI analyses from whole genomes indicated that strain 115 is closest to *P. saccharolyticum* (Fig. 4).

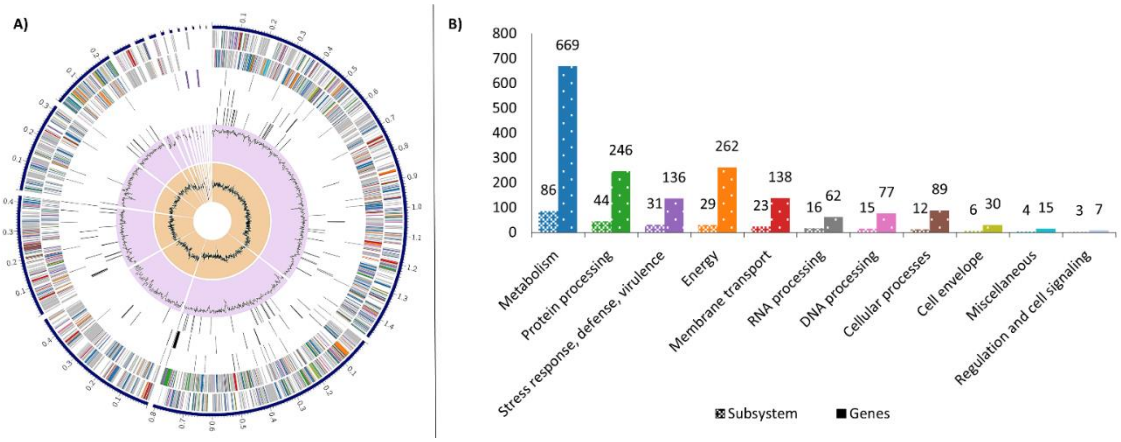


Figure 1: A) Circular graphical display of the distribution of the genome annotations from *Pseudochrobactrum saccharolyticum* strain 115 including, from outer to inner rings, the contigs, CDS on the forward strand, CDS on the reverse strand, RNA genes, CDS with homology to known antimicrobial resistance genes, CDS with homology to known virulence factors, GC content and GC skew. B) The colors of the CDS on the forward and reverse strand indicate the subsystem to which these genes belong.

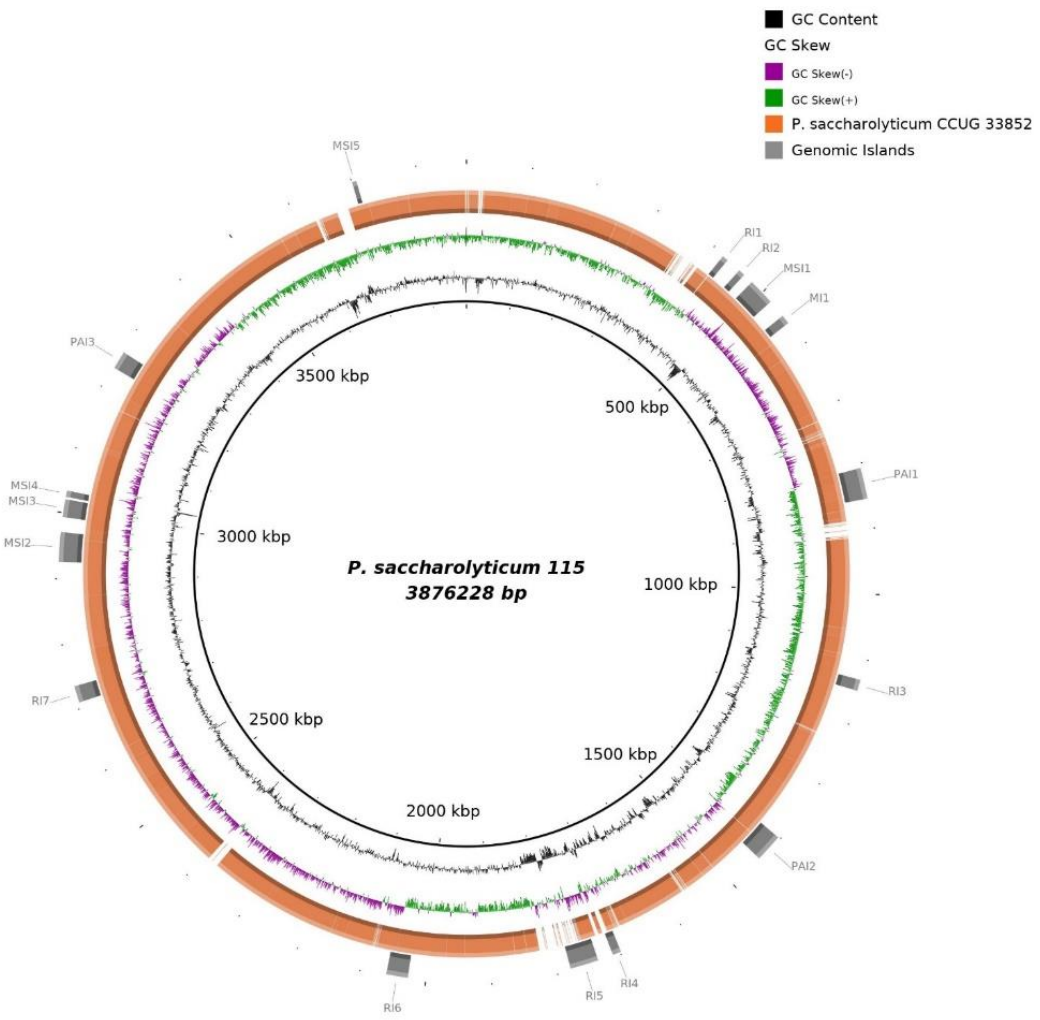


Figure 2: Circular representation showing location of genomic islands in the genomes of *Pseudochrobactrum saccharolyticum* strains. The type strain was aligned using strain 115 as a reference. The figure represents the coding sequences (CDS), Metabolism islands (MI), Pathogenicity islands (PAI), Antimicrobial resistance islands (RI), and Miscellaneous islands (MSI), with functions related to both metabolism, pathogenicity, and antimicrobial resistance.

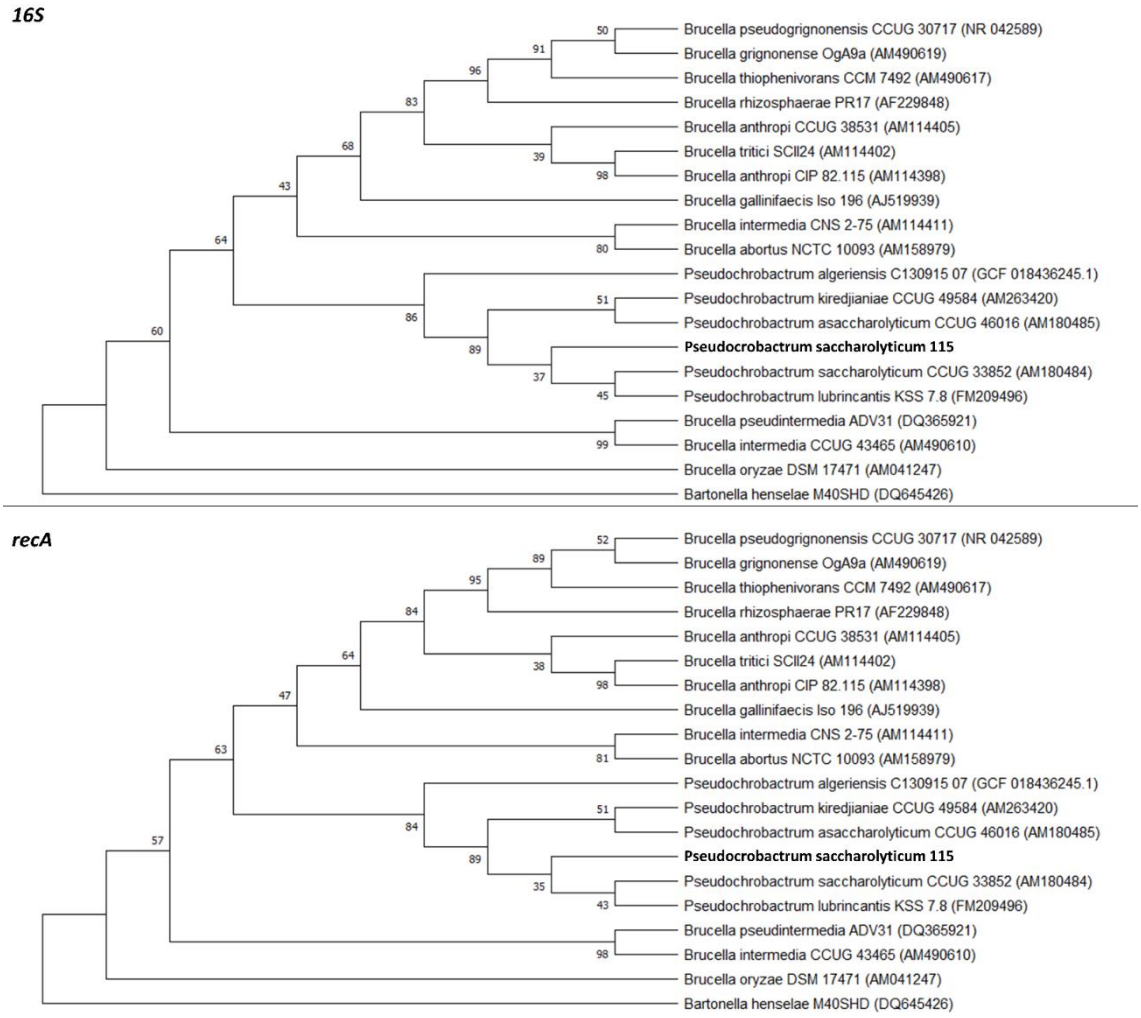


Figure 3: Phylogenetic analysis of *16S* rRNA and *recA* genes of representatives from the *Brucellaceae* family and *Bartonella* spp. strains. Trees were build using neighbor-joining reconstruction method in MEGA X v.11. Bootstrap values are represented in the branches. *Bartonella henselae* was used to root the tree.

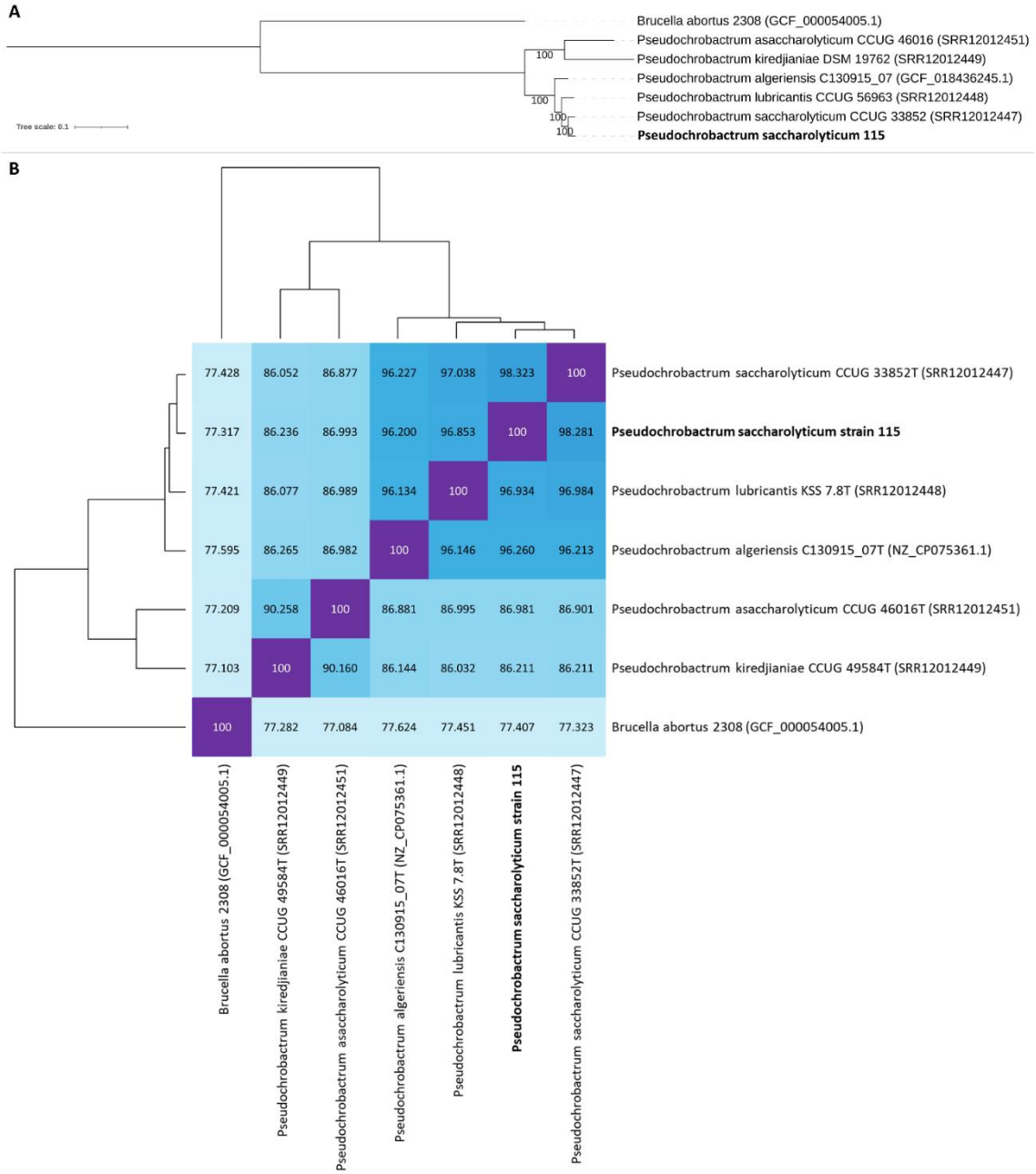


Figure 4: A) Phylogenomic analysis of *Pseudochrobactrum* spp. type strains and strain 115 by single nucleotide polymorphisms identified in whole genome sequences using Northern Arizona SNP Pipeline (NASP). Tree was build using Maximum-likelihood reconstruction in IQ-TREE. Bootstrap values are represented in the branches, with *Brucella abortus* 2308 used to root the tree. B) Average Nucleotide Identity (ANI) analysis performed in FastANI of *Pseudochrobactrum* spp. type strains and strain 115, using *Brucella abortus* 2308 to root the tree.

## DISCUSSION

*Pseudochrobactrum saccharolyticum* strain 115 was isolated from a bovine lymph node and initially misidentified as *Brucella* spp. Such a mistaken identification can have significant consequences, especially in regions whose brucellosis control is in advanced stages or in regions where the disease has already been eradicated, as in some European countries. In these regions, misdiagnosis of animal brucellosis can result in high costs involving epidemiological traceback studies and unnecessary culling of animals and may result in the temporary loss of brucellosis-free status and create barriers to national and international trade of animals and their products.

The initial identification as *Brucella* spp.*,* based on the phenotypic analysis and positive *Brucella* genus-specific PCR (BAILY; KRAHN; DRASAR; STOKER, 1992), was supported by the fact that it was identified from a cow. However, the negative result with the highly specific AMOS PCR, anchored on the IS711 insertion sequence that is only found in *Brucella* spp.*,* led us to doubt this identification. MALDI-TOF mass spectrometry is becoming an important tool in the identification of bacterial pathogens, and recently both the Bruker and bioMérieux systems have new spectral databases that allow for accurate identification of *Brucella* (MESUREUR; AREND; CELLIÈRE; COURAULT *et al.*, 2018). A MALDI-TOF platform also identified strain 115 as *Pseudochrobactrum* spp.*,* however it classified it as a closely related species, *P. asaccharolyticum.* In many countries, access to MALDI-TOF is still limited, but when available, the solvent inactivation step allows for safe handling of suspect *Brucella* spp. isolates on this instrument (MESUREUR; RANALDI; MONNIN; GIRARD *et al.*, 2016).

The *16S* and *recA* gene sequences proved to be useful in classifying the isolate at the genus level and provide more robust results for building phylogenetic trees (KÄMPFER; ROSSELLÓ-MORA; SCHOLZ; WELINDER-OLSSON *et al.*, 2006). Nonetheless, since the sequencing of few genes can also be an expensive and laborious method at a large-scale, the sequencing of *16S* and *recA* genes for diagnostic confirmation would be more feasible in brucellosis-free regions, where the identification of positive animals is rare and its impact at the population level is high. This method would be recommended in regions endemic for brucellosis only when atypical results are obtained, such as a positive amplification in *Brucella* genus-specific PCR (BAILY; KRAHN; DRASAR; STOKER, 1992) and negative reaction in AMOS-PCR (BRICKER; HALLING, 1994). The analysis of the strain 115 genome showed

that the result in *Brucella* genus-specific PCR was nonspecific for closely related *Brucellaceae* genera.

At the species level, the identification of strain 115 was only possible from whole genome analyses. SNPs and ANI analyses allowed the identification of strain 115 as *P. saccharolyticum* (ANI = 98.32%); however, as the genus *Pseudochrobactrum* is constantly expanding (LOPERENA-BARBER; KHAMES; LECLERCQ; ZYGMUNT *et al.*, 2022) and has very closely related species, with *P. algeriensis*, *P. lubricantis,* and *P. saccharolyticum* having ANI values greater than 95–96% (LOPERENA-BARBER; KHAMES; LECLERCQ; ZYGMUNT *et al.*, 2022). The cut-off point for species adopted for species classification followed PALMER; STEENKAMP; BLOM; HEDLUND *et al.* (2020) but it is important to note that species-level classification is limited by the current literature and taxonomic knowledge, both which continue to expand.

Considering that the *Pseudochrobactrum* genus has been isolated in different regions of the world such as Europe (ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.*, 2020), Africa (LOPERENA-BARBER; KHAMES; LECLERCQ; ZYGMUNT *et al.*, 2022) and Americas in the present study, suspect samples should be characterized beyond simple diagnostics. The ability to analyze and compare bacterial genomes has allowed unprecedented resolution to identify mechanisms of pathogens hitherto poorly understood, such as host-parasite relationships, virulence factors, and antimicrobial resistance, as well as genetic determinants involved in various metabolic functions. Genomic characterization of strain 115 identified its main subsystems, with the metabolism and protein processing responsible for more than 50% of the set of genes that jointly implement a specific biological process or structural complex in this microorganism. In the BRIG analysis, regions present only in the genome of the Brazilian strain were detected, differentiating it from the European strain *P. saccharolyticum* CCUG 33852$^{T}$. Some of these regions were classified as genomic islands in GIPSy analysis and demonstrate the genome plasticity of this species in acquiring large blocks of sequences by horizontal gene transfer, and, thus, showing the need for more studies to elucidate the impact of these events.

The addition of a new *P. saccharolyticum* genome is crucial for the *Brucellaceae* family, considering that the identification and isolation of *Pseudochrobactrum* spp. is unusual and is likely overlooked due to misidentification since conventional biochemical tests are not able to distinguish this genus from *Brucella* spp., as they share many characteristics in common (KÄMPFER; WOHLGEMUTH; SCHOLZ, 2014). Thus, reports on the identification of

*Pseudochrobactrum* spp. are limited to the articles on the first description of the species (KÄMPFER; HUBER; LODDERS; WARFOLOMEOW *et al.*, 2009; KÄMPFER; ROSSELLÓ-MORA; SCHOLZ; WELINDER-OLSSON *et al.*, 2006; KÄMPFER; SCHOLZ; HUBER; THUMMES *et al.*, 2007; LOPERENA-BARBER; KHAMES; LECLERCQ; ZYGMUNT *et al.*, 2022) and a recent phylogenomic study, in which some strains isolated from ruminants were sequenced (ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.*, 2020). The first *Pseudochrobactrum* spp. genome was deposited in 2013, with little epidemiological information available, hindering a better understanding of this microorganism. Nothing is known of *Pseudochrobactrum* spp. in Latin America as this is the first record.

Beyond the limited knowledge about the genetic repertoire of *Pseudochrobactrum* spp., another aspect that remains poorly studied in this genus is the susceptibility to antimicrobials. In the *Brucellaceae* family, “traditional” *Brucella* spp. strains are usually susceptible to a wide range of antimicrobial agents (SCHOLZ; BANAI; CLOECKAERT; KÄMPFER *et al.*, 2015). However, some multidrug-resistant strains have been recently described BARBOSA PAULETTI; STYNEN; MOL; DORNELES *et al.* (2015), while *Ochrobactrum* spp. (recently reclassified as *Brucella* spp.) is frequently resistant to several classes of drugs (TEYSSIER; MARCHANDIN; JEAN-PIERRE; DIEGO *et al.*, 2005). In this study, it was observed that high concentrations of rifampicin, streptomycin, and sulfamethoxazole + trimethoprim were necessary to inhibit the growth of *P. saccharolyticum* strain 115, suggesting that there may be some multidrug resistance. The potential for multidrug resistance is supported by the presence of products predicted in the subsystems identified in the PATRIC platform (Appendix A), e.g. Aminoglycoside N (6') – an acetyltransferase with potential to cause streptomycin resistance, in addition to the presence of drug efflux systems, such as RND type and MSF type, which can lead to multidrug resistance. The determination of breakpoints for the main antimicrobials used to treat human infections and the massive susceptibility test are crucial for a broad characterization of *Pseudochrobactrum* spp., since the isolation of these bacteria from both humans (KÄMPFER; ROSSELLÓ-MORA; SCHOLZ; WELINDER-OLSSON *et al.*, 2006) and animals (ASHFORD; MUCHOWSKI; KOYLASS; SCHOLZ *et al.*, 2020; LOPERENA-BARBER; KHAMES; LECLERCQ; ZYGMUNT *et al.*, 2022) along with the results of the present study suggests that this microorganism may be a zoonotic pathogen. Furthermore, it is important to consider its potential to donate mobile genetic elements to other bacteria by horizontal gene transfer, including virulence and antimicrobial resistance genes, which were

observed in the genomic island analysis and were present in strain 115 but absent in *P. saccharolyticum* type strain CCUG 33852$^{T}$.

This study reported the isolation and whole-genome sequencing of the first strain of *P. saccharolyticum* strain identified in Latin America. Furthermore, our findings indicated that the sequences of *16S* and *recA* alone could differentiate *Brucella* spp. and *Pseudochrobactrum* spp. strains, which would be especially important to exclude *Brucella* spp. false-positive results.

## SUPPLEMENTARY FILES

https://drive.google.com/drive/folders/1vOtahFFRv7pk0kWVbLRhQkvFIDeOEzKy?usp=drive_link

Appendix A: Matrix-assisted laser desorption ionization (MALDI)-time of flight (TOF) mass spectrometry (MS) spectrum analysis of *Pseudochrobactrum saccharolyticum* strain 115 using Biotyper from Bruker Daltonics.

Appendix B: Complete overview of the subsystems from *Pseudochrobactrum saccharolyticum* strain 115 provided by PATRIC.

Appendix C: Genomic islands and products identified in the analysis between the coding sequences of query (*P. saccharolyticum* strain 115) and subject genome (B. abortus strain 2308) generated by GIPSy.

Appendix D: Mauve analysis comparing the synteny of *Pseudochrobactrum saccharolyticum* strain 115 and *P. saccharolyticum* type strain CCUG 33852T.

## ACKNOWLEDGMENTS

We would like to acknowledge Marcus Vinicius Canário Viana, Raquel Costa Neia, and Saulo Britto da Silva for their technical support. This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq [423557/2016-5], Fundação de Amparo à Pesquisa do Estado de Minas Gerais – Fapemig [00283/2016], Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes and Network of Omics Sciences (RECOM). DO'C received a travel grant from CAPES/PrInt/UFLA [52/2019]. VBIC/CNR *Brucella* was supported by INSERM, Université de Montpellier and Santé Public France. DJJP was sponsored by the Wellcome Trust.

## REFERENCES

ALIKHAN, N. F.; PETTY, N. K.; ZAKOUR, N. L. B.; BEATSON, S. A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. **BMC Genom.**, 12, n. 1, p. 1-10, Aug 8 2011.

ALTON, G. G.; JONES, L. M.; ANGUS, R. D.; VERGER, J. M. **Techniques for the brucellosis laboratory**. Institut National de la recherche Agronomique (INRA), 1988. 2738000428.

ANDREWS, S.; KRUEGER, F.; SECONDS-PICHON, A.; BIGGINS, F. *et al.* **A quality control tool for high throughput sequence data**. Babraham Bioinformatics, 2010.

ASHFORD, R. T.; MUCHOWSKI, J.; KOYLASS, M.; SCHOLZ, H. C. *et al.* Application of Whole Genome Sequencing and Pan-Family Multi-Locus Sequence Analysis to Characterize Relationships Within the Family *Brucellaceae*. **Front. Microbiol.**, 11, p. 1329, 2020.

ASSIS, G. B. N.; PEREIRA, F. L.; ZEGARRA, A. U.; TAVARES, G. C. *et al.* Use of MALDI-TOF Mass Spectrometry for the Fast Identification of Gram-Positive Fish Pathogens. **Front. Microbiol.**, 8, n. 1492, p. 1492, 2017-August-09 2017.

BAILY, G. G.; KRAHN, J. B.; DRASAR, B. S.; STOKER, N. G. Detection of *Brucella melitensis* and *Brucella abortus* by DNA amplification. **Trop. Med. Int. Health**, 95, n. 4, p. 271-275, Aug 1992.

BANKEVICH, A.; NURK, S.; ANTIPOV, D.; GUREVICH, A. A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **J Comput Biol**. 19, n. 5, p. 455-477, 2012.

BARBOSA PAULETTI, R.; STYNEN, A. P. R.; MOL, J. P. S.; DORNELES, E. M. S. *et al.* Reduced Susceptibility to Rifampicin and Resistance to Multiple Antimicrobial Agents among *Brucella abortus* Isolates from Cattle in Brazil. **PLoS One**, 10, n. 7, p. e0132532, 2015.

BOSI, E.; DONATI, B.; GALARDINI, M.; BRUNETTI, S. *et al.* MeDuSa: a multi-draft based scaffolder. **Bioinformatics**, 31, n. 15, p. 2443-2451, Aug 1 2015.

BRICKER, B. J.; HALLING, S. M. Differentiation of *Brucella abortus* bv. 1, 2, and 4, *Brucella melitensis*, *Brucella ovis*, and *Brucella suis* bv. 1 by PCR. **J. Clin. Microbiol.**, 32, n. 11, p. 2660-2666, Nov 1994.

DARLING, A. C. E.; MAU, B.; BLATTNER, F. R.; PERNA, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. **Genome Res.**, 14, n. 7, p. 1394-1403, Jul 2004.

GU, Z.; EILS, R.; SCHLESNER, M. J. B. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. 32, n. 18, p. 2847-2849, 2016.

GUIZELINI, D.; RAITTZ, R. T.; CRUZ, L. M.; SOUZA, E. M. *et al.* GFinisher: a new strategy to refine and finish bacterial genome assemblies. **Sci. Rep.**, 6, p. 34963, 2016.

HÖRDT, A.; LÓPEZ, M. G.; MEIER-KOLTHOFF, J. P.; SCHLEUNING, M. *et al.* Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of Alphaproteobacteria. **Front. Microbiol.** 11, p. 468, 2020.

JAIN, C.; RODRIGUEZ-R, L. M.; PHILLIPPY, A. M.; KONSTANTINIDIS, K. T. *et al.* High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. **Nature Communications**. 9, n. 1, p. 1-8, 2018.

KÄMPFER, P.; HUBER, B.; LODDERS, N.; WARFOLOMEOW, I. *et al. Pseudochrobactrum lubricantis* sp. nov., isolated from a metal-working fluid. **Int. J. Syst. Evol. Microbiol.**, 59, n. Pt 10, p. 2464-2467, Oct 2009.

KÄMPFER, P.; ROSSELLÓ-MORA, R.; SCHOLZ, H. C.; WELINDER-OLSSON, C. *et al.* Description of *Pseudochrobactrum* gen. nov., with the two species *Pseudochrobactrum asaccharolyticum* sp. nov. and *Pseudochrobactrum saccharolyticum* sp. nov. **Int. J. Syst. Evol. Microbiol.**, 56, n. Pt 8, p. 1823-1829, Aug 2006.

KÄMPFER, P.; SCHOLZ, H.; HUBER, B.; THUMMES, K. *et al.* Description of *Pseudochrobactrum kiredjianiae* sp. nov. **Int. J. Syst. Evol. Microbiol.**, 57, n. Pt 4, p. 755-760, Apr 2007.

KÄMPFER, P.; WOHLGEMUTH, S.; SCHOLZ, H. The Family *Brucellaceae*. *In*: ROSENBERG, E.;DELONG, E. F.*, et al* (Ed.). **The Prokaryotes: Alphaproteobacteria and Betaproteobacteria**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. p. 155-178.

KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C. *et al.* MEGA X: molecular evolutionary genetics analysis across computing platforms. **Mol. Biol. Evol.**, 35, n. 6, p. 1547-1549, 2018.

LETUNIC, I.; BORK, P. J. B. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. **Bioinformatics**. 23, n. 1, p. 127-128, 2007.

LOPERENA-BARBER, M.; KHAMES, M.; LECLERCQ, S. O.; ZYGMUNT, M. S. *et al. Pseudochrobactrum algeriensis* sp. nov., isolated from lymph nodes of Algerian cattle. **Int J Syst Evol Microbiol**. 72, n. 2, p. 005223, 2022.

MARGULIES, M.; EGHOLM, M.; ALTMAN, W. E.; ATTIYA, S. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, 437, n. 7057, p. 376-380, 2005.

MESUREUR, J.; AREND, S.; CELLIÈRE, B.; COURAULT, P. *et al.* A MALDI-TOF MS database with broad genus coverage for species-level identification of *Brucella*. **PLoS Negl Trop Dis**, 12, n. 10, p. e0006874, 2018.

MESUREUR, J.; RANALDI, S.; MONNIN, V.; GIRARD, V. *et al.* A simple and safe protocol for preparing *Brucella* samples for matrix-assisted laser desorption ionization–time of flight mass spectrometry analysis. **J. Clin. Microbiol.**, 54, n. 2, p. 449-452, 2016.

MORENO, E.; BLASCO, J. M.; LETESSON, J. J.; GORVEL, J. P. *et al.* Pathogenicity and Its Implications in Taxonomy: The *Brucella* and *Ochrobactrum* Case. **Pathogens.** 11, n. 3, p. 377, 2022.

NGUYEN, L.-T.; SCHMIDT, H. A.; VON HAESELER, A.; MINH, B. Q. J. M. b. *et al.* IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. **Mol Biol Evol**. 32, n. 1, p. 268-274, 2015.

OREN, A.; GARRITY, G. List of new names and new combinations previously effectively, but not validly, published. **Int J Syst Evol Microbiol**. 70, n. 7, p. 4043-4049, 2020.

PALMER, M.; STEENKAMP, E. T.; BLOM, J.; HEDLUND, B. P. *et al.* All ANIs are not created equal: implications for prokaryotic species boundaries and integration of ANIs into polyphasic taxonomy. **Int J Syst Evol Microbiol.** 70, n. 4, p. 2937-2948, 2020.

SAHL, J. W.; LEMMER, D.; TRAVIS, J.; SCHUPP, J. M. *et al.* The Northern Arizona SNP Pipeline (NASP): accurate, flexible, and rapid identification of SNPs in WGS datasets. **bioRxiv.** p. 037267, 2016.

SCHOLZ, H. C.; BANAI, M.; CLOECKAERT, A.; KÄMPFER, P. *et al. Brucella*. **Bergey's Manual of Systematics of Archaea and Bacteria**, p. 1-38, 2015.

SCHOLZ, H. C.; TOMASO, H.; DAHOUK, S. A.; WITTE, A. *et al.* Genotyping of *Ochrobactrum anthropi* by *recA*-based comparative sequence, PCR-RFLP, and *16S rRNA* gene analysis. **FEMS Microbiol. Lett.**, 257, n. 1, p. 7-16, 2006.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, 30, n. 14, p. 2068-2069, 2014.

SOARES, S. C.; GEYIK, H.; RAMOS, R. T. J.; DE SÁ, P. H. C. G. *et al.* GIPSy: genomic island prediction software. **J. Biotechnol.**, 232, p. 2-11, 2016.

TEYSSIER, C.; MARCHANDIN, H.; JEAN-PIERRE, H.; DIEGO, I. *et al.* Molecular and phenotypic features for identification of the opportunistic pathogens *Ochrobactrum* spp. **Int. J. Med. Microbiol.**, 54, n. 10, p. 945-953, 2005.

VILA, A.; PAGELLA, H.; BELLO, G. V.; VICENTE, A. Brucella suis bacteremia misidentified as *Ochrobactrum anthropi* by the VITEK 2 system. **J. Infect. Dev. Ctries.**, 10, n. 04, p. 432-436, 2016.

WARNES, M. G. R.; BOLKER, B.; BONEBAKKER, L.; GENTLEMAN, R. *et al.* Package 'gplots'. 2016.

WHATMORE, A. M.; FOSTER, J. T. Emerging diversity and ongoing expansion of the genus *Brucella*. **Infection, Genetics and Evolution**, 92, p. 104865, 2021/08/01/ 2021.

WICKHAM, H. J. J. o. s. s. Reshaping data with the reshape package. 21, p. 1-20, 2007.

## FINAL CONSIDERATIONS

Overall, the results of this thesis pointed to the large number of approaches that whole genome sequencing can be applied, especially when associated with the isolate metadata. It was possible to observe that the identification of the strains at the epidemiological level, such as year of strain isolation, state of origin and host, together with the microbiological characterization, such as biovar and resistance to antimicrobials, were essential for the correct interpretation of the bioinformatics analyzes carried out. The *B. abortus* strains isolated and sequenced in Brazil showed a high genetic clonality, especially within the same biovar classification. The most indicated method for the surveillance of this pathogen in Brazil was the MLVA16, except for Santa Catarina where the cgMLST would be more suitable. It was also possible to identify that the antimicrobial resistance mechanisms already described in the literature do not seem to explain the drug resistance phenotype in the Brazilian isolates, and this answer may be in the regulatory genes, which are not conventionally reported as responsible for this phenotype in other pathogens. The Brazilian genomes added 34 genes to the genetic repertoire of *B. abortus* and exhibited an excellent quality when submitted to curatorship analysis prior to the performance of the pan-genome analysis. Finally, the isolation of a genus previously misclassified as *Brucella* spp. was reported for the first time in the country and thanks to the application of whole genome sequencing technologies it was possible to correctly identify the microorganism and raise important discussions about the use of this tool to avoid false-positive diagnoses of brucellosis. Genomic surveillance constitutes the future for the pathogen control and eradication programs and should always be used in association with classical conventional investigation methods to explore the full potential of its application and turn its responses into practical prevention measures of both human and veterinary diseases.