

Comparative genomics of *Brucella suis* discriminates different biovars

Kumaragurubaran Karthik¹, Subbaiyan Anbazhagan² and Seeralan Manoharan^{3*}

¹Veterinary College and Research Institute, Udumalpet, Tamil Nadu Veterinary and Animal Sciences University, Chennai, India

²ICMR-National Animal Resource Facility for Biomedical Research, Hyderabad, India

³Vaccine Research Centre- Bacterial Vaccines, Centre for Animal Health Studies, Tamil Nadu Veterinary and Animal Sciences University, Chennai, India

KARTHIK, K., S. ANBAZHAGAN, S. MANOHARAN: Comparative genomics of *Brucella suis* discriminates different biovars. Vet. arhiv 94, 285-296, 2024.

ABSTRACT

Brucella suis TANUVAS_1 was isolated from a case of porcine abortion and whole genome sequenced. In the present study, comparative genomics of *B. suis* strains (n=33), including the newly sequenced strain TANUVAS_1 was carried out for accurate discrimination of the biovars. To alleviate the problems associated with conventional biovar identification techniques, a whole genome-based approach was used in the present study. The whole genome phylogeny of *B. suis* clearly differentiated strains on the basis of biovars. A total of 43 virulence genes were predicted, of which the *btpA* gene was present in only 14 strains, whereas the *wbkA* gene was absent in 19 strains. Multi locus sequence typing (MLST) analysis showed 5 different sequence types (14, 15, 16, 17 and 19) and the maximum number of strains belonged to ST14. There were 6 different STs assigned on the basis of core genome MLST, and the maximum number of strains belonged to ST50. Pangenome analysis of the strains showed that a total of 1948 genes (30.47%) were core genes and 4445 genes (69.53%) were accessory genes, suggesting that the *B. suis* genome has an open pan genomic structure. SNP analysis revealed 17 biovar 1 specific SNPs, 2142 biovar 2 specific SNPs, 673 biovar 3 specific SNPs, 714 biovar 4 specific SNPs and 2308 biovar 5 specific SNPs. The results show that cgMLST can be used for accurate discrimination of strains on the basis of biovars. Similarly, biovar specific SNPs can be employed for developing molecular assays that can discriminate biovars.

Key words: *Brucella suis*; comparative genomics; biovar specific genes; biovar specific SNPs; virulence genes

Introduction

Brucella suis is a potentially zoonotic agent that can affect several animals, such as domestic pigs, wild pigs, hares, caribou wolves, bears, reindeers, foxes, cattle, horses, sheep, yak, and rodents. *B. suis* has 5 biovars (biovars 1 to 5) and biovars 1 and 3 are reported to be pathogenic to humans (GARIN-

BASTUJI and HARS, 2001). *B. suis* biovar 2 has lower infection potential in healthy humans

(ALLEN et al., 1998) but has recently been isolated from dogs (GIRAULT et al., 2023). *B. suis* biovars differ according to geographical region, host, and the severity of infection (OIE, 2013).

*Corresponding author:

Seeralan Manoharan, Vaccine Research Centre- Bacterial Vaccines, Centre for Animal Health Studies, Tamil Nadu Veterinary and Animal Sciences University, Chennai-600051, India, e-mail: manoulagai@gmail.com

Wild boars and hares can act as reservoirs of *B. suis*, and can transmit the infection to domestic pigs (GRE'GOIRE et al., 2012). Hence, it is essential to differentiate biovars so that prevention and control measures can be formulated.

B. suis biovars can be classified using conventional microbiological methods, but these methods are time-consuming and can be hazardous to the handlers. Conventional methods can sometimes lead to misinterpretation of strains, and there is also a risk for the handlers since the organism is zoonotic. Molecular methods such as restriction fragment length polymorphism PCR (RFLP), multilocus sequence typing (MLST), multiple locus variable number of tandem repeat analysis (MLVA), variable number tandem repeat (VNTR), SNP typing, and matrix-assisted laser desorption ionization-time of flight mass spectrometry analysis (MALDI-TOF) can be used for discrimination of strains (AL DAHOUK et al., 2005; LE FLÈCHE et al., 2006; WHATMORE et al., 2007; MESUREUR et al., 2016). Suis-ladder PCR can be used for differentiation of *B. suis* biovars, but requires two steps of multiplex PCR, and the discriminatory power of the assay is lower compared to whole genome-based assays (LÓPEZ-GOÑI et al., 2011).

The whole genome sequence-based approaches, such as core genome MLST, pan-genome analysis and whole genome SNP analysis, are powerful tools for discrimination of strains, and hence can be used for epidemiological surveillance (GIRAULT et al., 2022). Whole genome phylogeny and pan-genome analysis can discriminate various strains of *Brucella melitensis* (KARTHIK et al., 2021). SNP based phylogeny has already been shown to discriminate biovars of *B. suis* (SANKARA SUBRAMANIAN et al., 2015). It was also reported earlier that *B. suis* had inter species diversity, based on whole genome phylogeny, but these methods require whole genome sequencing of the strains (FOSTER et al., 2009). Since whole genome sequence based comparative genomics are limited, especially for *B. suis*, the present study was undertaken to compare genome sequences from different regions. Methods such as whole genome-based pan-genome analysis,

MLST, cgMLST, SNPs were used in the present study to identify a suitable molecular method for discrimination of *B. suis* strains.

Materials and methods

Isolation and identification of B. suis. A total of 42 field samples from pigs (22 aborted foetal contents and 20 whole blood samples), suspected to have brucellosis, were received from a private farm in Vedaranyam, Nagapattinam district of Tamil Nadu state, India, during the month of March, 2017. The samples were cultured in *Brucella* selective broth (Himedia, India), and were then sub cultured on *Brucella* selective agar with 5 % horse serum as a supplement. Suspected colonies were gram stained and DNAs was extracted from these colonies using a DNeasy Blood and Tissue Kit (Qiagen, USA). AMOS PCR was carried out for confirmation of *Brucella* spp. (BRICKER and HALLING, 1994).

Whole genome sequencing, assembly, and annotation. Genomic DNA from the *B. suis* TANUVAS_1 strain was extracted using the DNeasy Blood and Tissue Kit (Qiagen, USA) as per the manufacturer's protocol, and submitted for NGS to Gujarat Biotechnology Research Center, Department of Science & Technology, Gujarat. Illumina Miseq sequencing technology was used for generation of 2 x 150 bp paired-end reads. For testing the quality of the raw reads, the FastQC tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was used. For genome assembly, Unicycler version 0.4.8.0 was used, with default options which function as a SPA des-optimizer, and for annotation, Prokka 1.14.5 and the Rapid Annotations utilizing Subsystems Technology (RAST) annotation pipeline were used (AZIZ et al., 2008; TATUSOVA et al., 2016; WICK et al., 2017).

Genome data retrieval and whole genome phylogeny. The complete *B. suis* genome from NCBI was used for analysis. A total of 32 complete genome assemblies with respect to all the *B. suis* biovars were retrieved from NCBI, as of 11.09.2022 (Table 1).

Table 1. Metadata of the *B. suis* strains used in the study

Biovar	Strain	Assembly	Collection year	Place
unknown	2004000577	GCA_001715425.1	2003	USA: Massachusetts
unknown	Human/AR/ US/1981	GCA_000875695.1	1981	USA:Arkansas
unknown	QH05	GCA_002749475.1	2013	China:Qinghai Province
unknown	ZW043	GCA_000755105.1	unknown	Zimbabwe: Chiredzi
unknown	ZW046	GCA_000755085.1	unknown	Zimbabwe: Norton
1	1330	GCA_000007505.1	unknown	unknown
unknown	1330	GCA_000223195.1	unknown	unknown
1	19	GCA_001594165.1	1983	China
unknown	TANUVAS_1	-	2017	INDIA
unknown	ML9	GCA_003194855.1	2017	India:Meghalaya
1	VBI22	GCA_000236255.1	unknown	unknown
1	CVI_59	GCA_014884545.1	2003	Croatia
1	CVI_58	GCA_014884565.1	2003	Croatia
1	60	GCA_009218085.1	2003	Croatia
1	S2	GCA_000600055.1	1960	China
2	ATCC 23445	GCA_000018905.1	unknown	unknown
2	Bs364CITA	GCA_000698325.1	2008	Spain
2	Bs396CITA	GCA_000698345.1	2009	Spain
2	PT09172	GCA_000698285.1	2009	Portugal
2	PT09143	GCA_000698245.1	2009	Portugal
2	Bs143CITA	GCA_000699025.1	2005	Spain
2	CVI_213	GCA_014884405.1	2017	Slovenia
2	CVI_76	GCA_014884465.1	2004	Croatia
2	CVI_50	GCA_014884585.1	2000	Croatia
2	CVI_196	GCA_014884425.1	2015	Croatia
2	CVI_105	GCA_014884445.1	2009	Croatia
2	BCB032	GCA_000292105.2	2008	China
3	CVI_71	GCA_014884525.1	2004	Slovenia
3	686	GCA_000740255.1	2005	unknown
4	CVI_72	GCA_014884505.1	2004	Slovenia
4	BSP	GCA_000740435.1	unknown	unknown
5	CVI_73	GCA_014884485.1	2004	Slovenia
unknown	513UK	GCA_000740235.1	unknown	United Kingdom

Multiple sequence alignments of all the genomes were carried out in a Realphy server (<https://realphy.unibas.ch/realphy/>). The phylogenetic relatedness of all the genomes along with the newly sequenced genome was found using IQTREE software, and the GTR+G model was used to reconstruct the phylogenetic tree, with 1000 bootstrap values (TRIFINOPOULOS et al., 2016). The unrooted tree was rooted with a mid-root tree. Constructed phylogeny was visualized and annotated using the Interactive Tree of Life (iTOL v6) (LETUNIC and BORK, 2007).

Pan-genome analysis. The comparative pan-genomic study targeted all the complete *B. suis* genomes available in NCBI (n=32) and the newly sequenced Indian strain (n= 1). A Panaroo tool and Roary 13.0 was used for pan-genome analysis (PAGE et al., 2015; TONKIN-HILL et al., 2020). FriPan was used for visualization of the two-dimensional scaling and accessory genome dendrogram (<https://github.com/drpowell/FriPan>). PanGP software was used to explore the pan-genome profile and unique gene status of *B. suis* (ZHAO et al., 2014). The pan-genome curve fitting was performed using Heaps law ($y=A \text{ pan}^x \text{ Bpan} + C \text{ pan}$) to identify the openness of the genome. Since there was a total of 33 genomes, PanGP recommended a distance guide algorithm for profiling pan-genome with roary binary output RTAB file (TETTELIN et al., 2005; RASKO et al., 2008).

Multilocus sequence typing analysis and virulence gene prediction. The allelic profile and sequence types (ST) of 33 *B. suis* genomes were identified using MLST, scanning genomes against PubMLST schemes (Galaxy Version 2.19.0). The available brucella pubMLST scheme and minimum DNA identity of 95% and DNA coverage of 10% were used for determining similar STs (SEEMANN, 2016a). cgMLST analysis was carried out in the pubMLST scheme to identify the closest cgST for the genomes. Genome assembly files were used for the cgMLST.

The ABRicate tool available in the Galaxy server (Galaxy Version 1.0.1) was used for screening of virulence genes in the genome assemblies with the VFDB database. For this, 80% DNA identity and coverage were used (SEEMANN, 2016b). The virulence gene presence-absence was represented

using the Displayr web tool (<https://app.displayr.com/>).

SNP Analysis. kSNP3.0 was used to do an SNP analysis that was independent of multiple alignments and the reference genome (GARDNER et al., 2015). Using the tool Kchooser included with the package, the ideal size of the nucleotide areas bordering the SNPs (kmer) was determined. The tool's SNP matrix file was used to create a maximum-parsimony tree and perform SNP analysis. Further whole genome alignment from REALPHY was used for the snp-distance matrix analysis in the galaxy server (SEEMANN, 2019).

Results

Genome of *B. suis* TANUVAS_1. Only one isolate was recovered from the 42 samples subjected to isolation and identification. There was amplification of 285 bp in the AMOS PCR corresponding to *B. suis*. The genome of the *B. suis* TANUVAS_1 strain had 29 contigs that were more than 1000 bp in size, and the actual genome size was calculated as 3.38Mb. The genome had 57.3% GC content, with N50 and L50 363417 and 4, respectively. A total number of 3293 CDS, 306 subsystems and 50 RNAs were predicted.

Whole genome phylogeny. Two major clades could be visualized on whole genome phylogeny, and there was clear distinction of the biovars of *B. suis*, except for BCB032, a biovar 2 strain clustering with the biovar 1 clade (Fig. 1).

Biovars of several strains, such as Human/AR/US/1981, 2004000577, ZW046, 1330, ZW043, QH05, TANUVAS_1 and ML9 clustering with biovar 1 clade, were not known. The whole genome phylogeny clusters the isolates of unknown biovars with known biovar groups. Therefore, the strain BCB032 grouping with a different biovar needs to be rechecked for its biovar. All biovar 2 strains in the present study belonged to Europe. None of the biovar 2 strains were isolated from humans. There was no clustering based on host and year of isolation. Strains CVI_71 and 686 corresponded to biovar 3, which forms a separate clade. Similarly, biovar 4 strains CVI_72 and BSP clustered together. CVI_73, a biovar 5 strain, was clustering with 513UK, the biovar of which was not known.

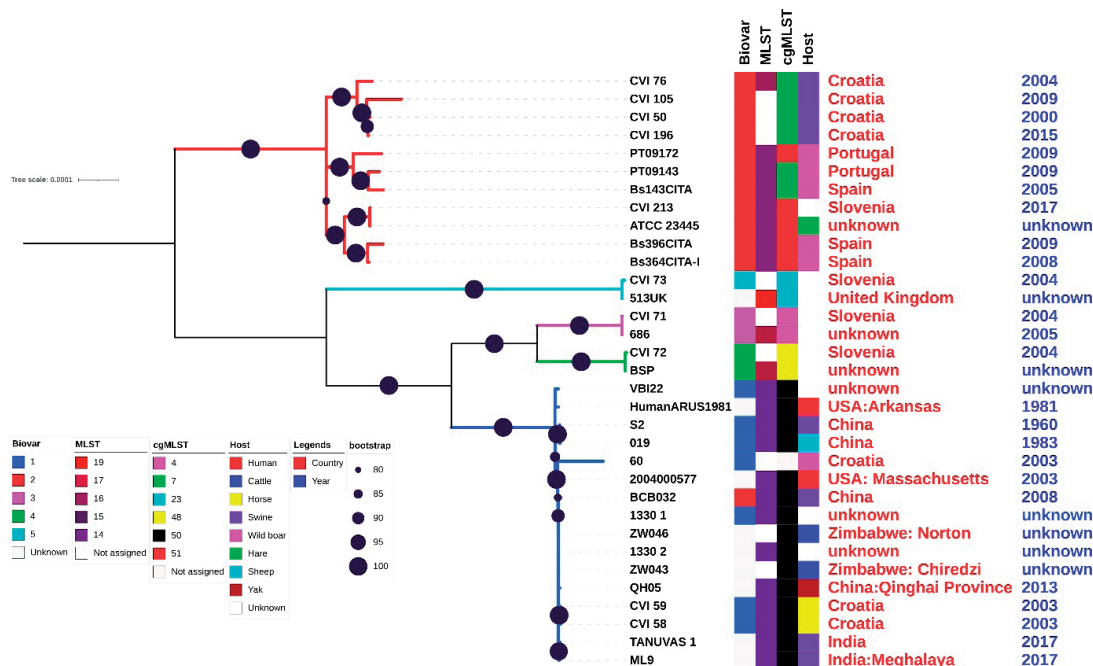


Fig. 1. Whole genome phylogeny of *Brucella suis*

Phylogeny with discrimination of biovars, MLST, cgMLST sequence types. Strain BCB032, a biovar 2 strain, is clustering with the biovar 1 clade

Pan-genome. A total of 6393 genes were predicted by pan-genome analysis, of which 1948 genes (30.47%) were core genes, and 4445 genes (69.53%) were accessory genes. Among the accessory genes, 557, 899 and 2989 were predicted to be soft-core genes, shell genes and cloud genes,

respectively. Pan-genome curve fitting analysis showed that the B_{pan} value was 0.68, indicating that the *B. suis* genome could be considered an open pan-genome. For every new genome addition to the pan-genome analysis, about 70 genes were added to the pan-genome (Fig. 2).

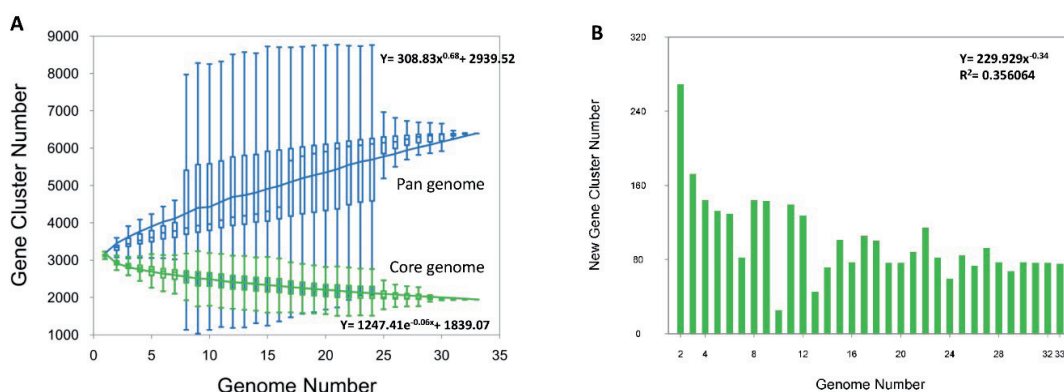


Fig. 2. Pan-genome curve of *B. suis*

A. Gene accumulation curves of the pan-genome (blue) and core-genome (green) of *B. suis*. B. The number of new genes added with an increase in the number of new *B. suis* genomes

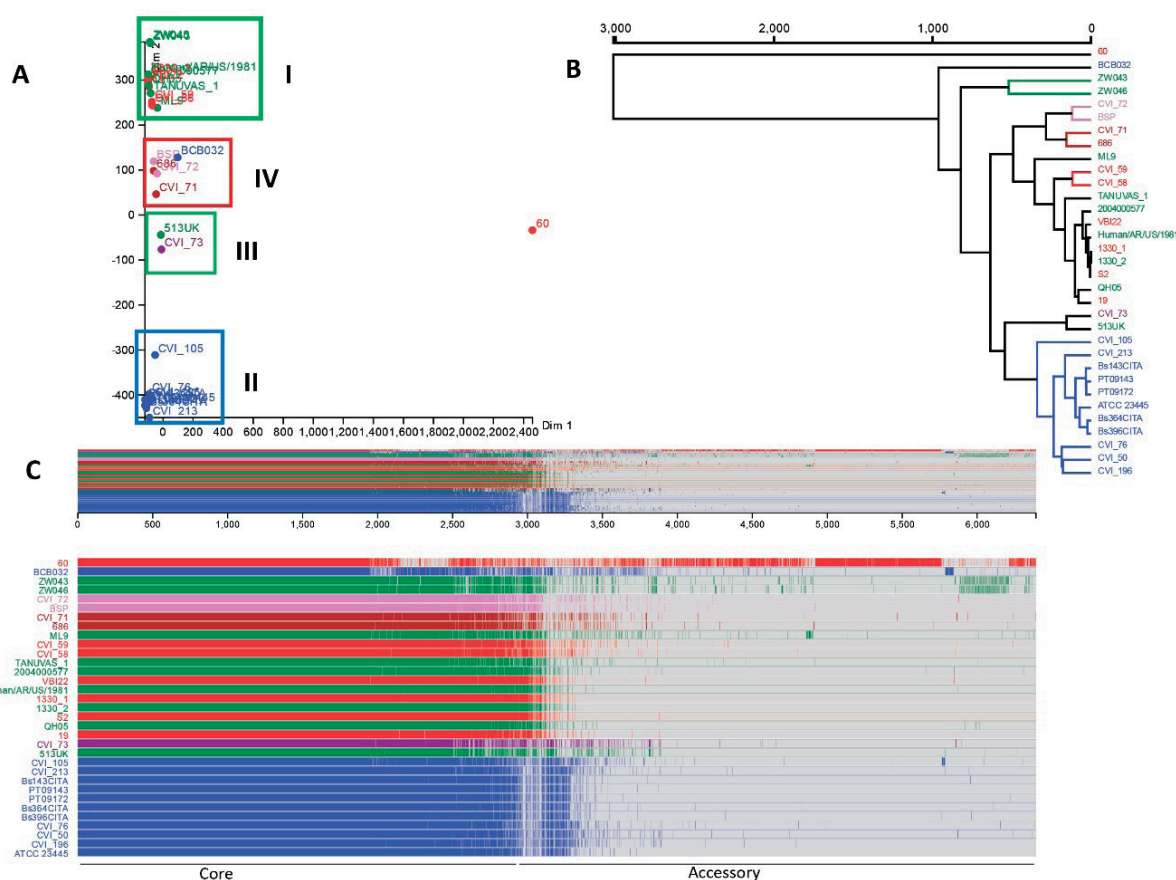


Fig. 3. Results of pan-genome analysis

A. Multidimensional scaling representing 4 clusters with different biovars. B. Pan-genome phylogeny with clusters marked with respective colors as in multidimensional scaling C. Schematic representation of core and accessory genes among the various genome analyzed

Multi-dimensional scaling had 4 major clusters and one single genome. Cluster I corresponded to biovar 1, cluster II belonged to biovar 2, cluster III belonged to biovar 5 and cluster IV had a genome from biovars 3, 4, and one biovar 2 genome. Strain 60 was unique since it did not cluster with any of the clusters in multi-dimensional scaling. Similarly, strain BCB032, a biovar 2 genome, was clustering with biovar 3 and 4 clusters (Fig.3).

Except for the genome of strains 60 and BCB032, which were out grouped from their respective biovar clades, pangenome phylogeny mimicked whole genome phylogeny.

Virulence genes. ABRicate predicted 43 virulence genes in the genome of *B. suis*. Among the 43 genes, only 14 genomes, namely: 513UK, ATCC_23445, BCB032, Bs143CITA, Bs364CITA-I, Bs396CITA, PT09172, PT09143, CVI_76, CVI_73, CVI_50, CVI_213, CVI_196 and CVI_105, were predicted with the *btpA* gene. All biovar 2 strains and two strains clustering in biovar 5 clade of whole genome phylogeny, had the *btpA* gene. The gene *wbkA* was absent in strain 19, while the *cgcs* gene was absent in the ZW043 strain. Indian strains TANUVAS_1 and ML9 were predicted with 42 virulence genes, and lacked the *btpA* gene (Fig. 4).

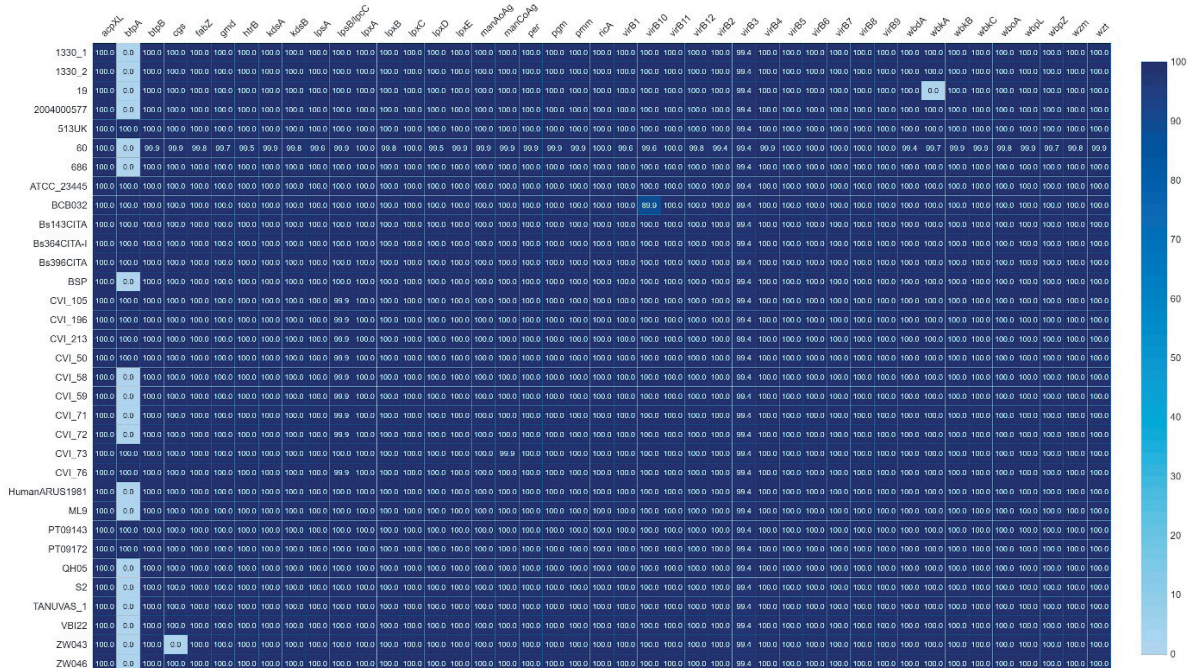


Fig. 4. *B. suis* virulence gene presence or absence chart

0 represents the absence of the gene, and 100 represents the complete presence of the gene, while values less than 100 indicate the presence of gene with less similarity when compared to the reference

MLST. A total of 5 STs (14, 15, 16, 17 and 19) were identified. A maximum of 13 strains belonged to ST14, 7 strains belonged to ST15, 2 strains (686 and BSP) belonged to ST17, and one strain each belonged to ST16 (strain CVI_76) and ST19 (strain 513UK). All biovar 1 strains belonged to ST14, while 7 biovar 2 strains belonged to ST15. The Indian strains TANUVAS_1 and ML9 belonged to ST14. Biovar 5 strain CVI_73 is a single locus (*cobQ*) variant of ST19. Similarly, strains CVI_71 and CVI_72 are single loci variants (*cobQ*) of ST17. Strain CVI_50 is a single loci variant (*dnaK*) of ST16, while CVI_196 and CVI_105 are double loci variants of ST16.

Except for genome 60, all the other genomes were assigned tocgMLST. All the genomes in biovar 1 cluster belonged to ST50. Biovar 2 strain

BCB032 genome also belonged to ST50. Genomes ATCC 23445, Bs364CITA, Bs396CITA, PT09172 and CVI_213 were assigned to ST51. Other biovar 2 strains, PT09143, Bs143CITA, CVI_76, CVI_50, CVI_196 and CVI_105, belonged to ST7. Biovar 3 strains CVI_71 and 686 belonged to ST4, while biovar 4 strains CVI_72 and BSP belonged to ST48. Strains CVI_73 and 513UK belonged to ST23 (Fig. 1).

SNP diversity. Reference free kSNP3 analysis showed a total of 10906 SNPs, of which 5536 were non-synonymous (NS) and 3255 were synonymous (S) SNPs. The NS/S ratio was 1.7, indicating diversifying population/positive selections. SNP analysis also showed biovar based grouping of the strains (Fig. 5).

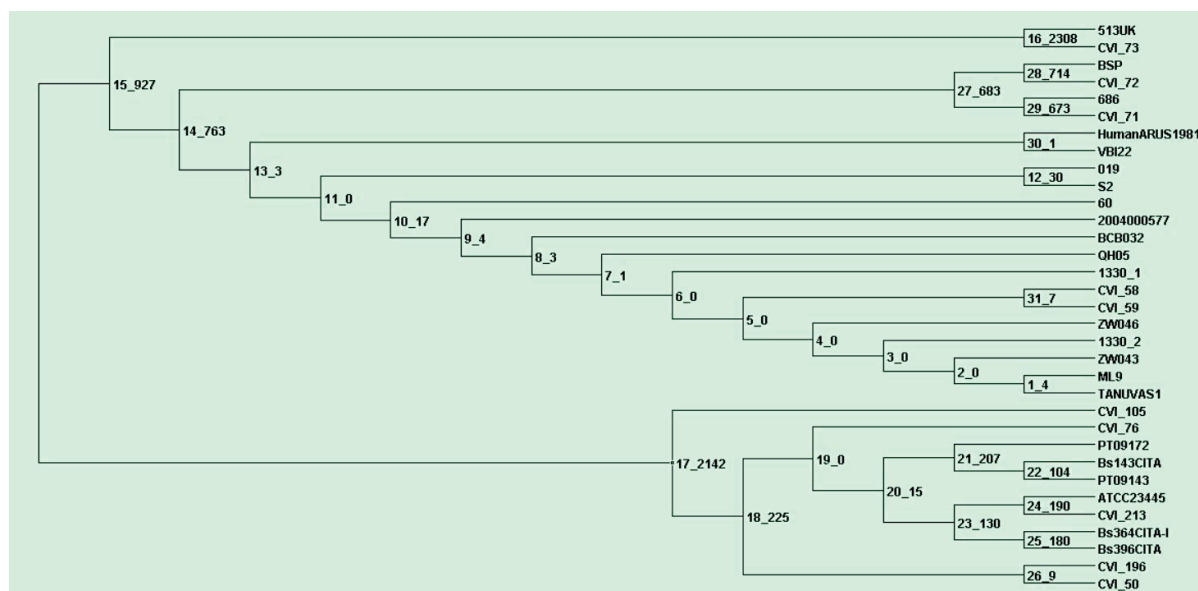


Fig. 5. SNP based phylogeny

Numbers after underscore (_) indicate the number of SNPs between strains

Strain 60 had a maximum of 315 strain specific SNPs among all the strains. A total of 17 biovar 1 (except VB122, HUMANARUS 1981, S2 and 019) specific SNPs (12 non-synonymous and 3 synonymous SNPs) were identified. Biovar 2 (except BCB032) specific 2142 SNPs (1162 non-synonymous and 626 synonymous SNPs) were identified. Biovar 2 specific non-synonymous SNPs were found in *virB2*, *virB4* and *virB5* virulence genes. A total of 673 biovar 3 specific SNPs were identified, of which 335 were non-synonymous, while 223 were synonymous SNPs. Similarly, 714 biovar 4 specific SNPs (356 non-synonymous SNPs and 203 synonymous SNPs) were identified. A maximum of 2308 SNPs (1201 non-synonymous SNPs and 650 synonymous SNPs) were identified as biovar 5 specific SNPs. A non-synonymous SNP of the *virB11* gene was found in biovars 2 and 5. Similarly, a non-synonymous SNP of *virB3* and *virB10* was found in the biovar 5 strains. Virulence gene *virB5* had non-synonymous SNP in biovar 3 and 5 strains. The gene responsible for intracellular survival, *cgs*, was found to have non-synonymous SNP in biovar strains 2, 3, 4, 5. Non-synonymous SNP was found in the phosphor-mannomutase

gene, *pmm* in biovars 2, 5 and strain PT09172. Similarly, non-synonymous SNPs were also found in GDP-mannose 4,6-dehydratase gene *gmd* in biovars 2, 4 and 5. Indian strains TANUVAS_1 and ML9 were differentiated by non-synonymous SNPs in 4 genes responsible for glycerol kinase GlpK, YafY family transcriptional regulator, ABC transporter permease, and FOF1 ATP synthase subunit beta.

Discussion

In the present study, whole genome phylogeny could clearly distinguish different biovars, except the BCB032 strain, which was contradictory to SANKARA SUBRAMANIAN et al. (2015), where whole genome phylogeny could not clearly differentiate the biovars. A recent study by PAPIĆ et al. (2023) also supports our results of biovar based distinction of WGS phylogeny. The genome used in both studies is not common. Biovars 2 and 5 were clearly distinct in SANKARA SUBRAMANIAN et al. (2015) which were similar to this study. Similarly, SANKARA SUBRAMANIAN et al. (2015) reported the BCB032 strain as biovar 1

but according to NCBI genome records, the strain BCB032 is biovar 2. Although strain 513UK clusters with biovar 5, and ZW043 and ZW046 with biovar 1, there is no clear record for the biovars in the NCBI database.

The *btpA* gene is involved in intracellular survival and unsettling the pathways of the host cells (CORONAS-SERNA et al., 2020). In the present study, 19 genomes lacked the *btpA* gene while it was present in biovars 2 and 5. Similarly, it was earlier reported that *btpA* was absent in *B. suis* 1330 (SALCEDO et al., 2013). Hence, further studies are required to understand the role of the *btpA* gene in biovar 2 and 5 strains. Similarly, the absence of *wbkA* in strain 19 and the *cgs* gene in strain ZW043 needs to be explored. The *wbkA* gene is responsible for the smoothness of the strain. It has been reported earlier that *wbkA* was disrupted in strain 19, and hence this may be the reason for the rough phenotype of the strain (WANG et al., 2016).

Strain CVI76 of biovar 2 was assigned ST16 by MLST, which was the least assigned ST among strains recovered from Europe, as most of the strains belonged to ST15 (DI SABATINO et al., 2017). WHATMORE et al. (2007) reported biovars 1, 2, 3, 4 and 5 were assigned to ST14, ST15/16, ST17, ST17/18 and ST19, respectively. Biovar 2 was assigned to ST7 and ST51 by cgMLST, while all other biovars were assigned to individual STs. Thus, cgMLST is more closely related to biovar classification and has better discrimination than MLST. Hence, cgMLST can be a better tool than MLST for the discrimination of strains, but cgMLST requires complete genome sequencing. The strain BCB032, a biovar 2, was found to be clustered with biovar 1 by whole genome phylogeny, pan-genome phylogeny, MLST and cgMLST. Hence, this strain needs further biochemical characterization to confirm its biovar.

Pan-genome analysis specifically for *B. suis* has not been reported previously, to the best of our knowledge. Pan-genome analysis showed that the size of the accessory genome was higher than the core genome, and new genes were added when new strains were added to the analysis. As the pan-genome analysis indicated that *B. suis* has an open

genome, there are chances that new strains will have novel genes (SANKARA SUBRAMANIAN et al., 2015). SNPs identified as specific for biovars can be used as molecular markers for typing the *B. suis* strains. A very recent study also identified 17 *Brucella* species specific SNP markers, and *B. suis* biovar specific SNP based primers were also designed for High-Resolution Melting PCR. The designed primers can differentiate different biovars of *B. suis* (GIRAULT et al., 2022). Similarly, genotype specific SNPs were identified for *B. melitensis*, and SNP-based phylogenetic analysis can be a tool for discrimination of strains (TAN et al., 2015). Thus, the present study shows that cgMLST and whole genome phylogeny can be used for discrimination of biovars, but these methods require complete genome sequencing. The biovar specific SNPs identified can be used for developing PCR-based assays that can easily discriminate biovars. PCR-based discrimination of biovars is an easy and cost-effective method compared to whole genome sequencing. Moreover, biovar identification by conventional methods requires handling live cultures, which presents a risk to operators of infection. Also, isolation of *Brucella* spp. is time consuming, which might be an obstacle for diagnosis, prevention, and control of the disease (AURICH et al., 2023). On the other hand, there may be a chance of misinterpretation of a biovar by a conventional method, as with the strain BCB032. Since only 32 complete genomes were available in NCBI, other genomes were not included for analysis and several more strains do not have any information regarding the biovars. Hence, a clear categorization of biovar specific genes cannot be claimed authoritatively. These lacunae are the impeding factors of this study. Hence, the genes predicted as biovar specific in the present study can be used in the future with a larger number of isolates for validation.

Conclusions

From an epidemiological point of view, it will be useful to know the circulating biovars in a particular geographical location. Hence, tools such as cgMLST or biovar specific SNPs identified in this study, can be used for accurate classification of

the species and biovars of the *Brucella* genus that are prevailing in a particular area.

Availability of supporting data

Genome sequence of *B. suis* TANUVAS_1 was submitted to NCBI SRA. Corresponding Bio Project number is PRJNA839908.

Author contributions

Kumaragurubaran Karthik and Subbaiyan Anbazhagan contributed equally to the work. Kumaragurubaran Karthik: conceptualization, software analysis, formal analysis, writing original draft; Subbaiyan Anbazhagan: data curation, software analysis, updating and reviewing manuscript; Seeralan Manoharan: *B. suis* isolation, identification, confirmation of *B. suis* TANUVAS_1, conceptualization and coordination of the work.

Declaration of Competing Interest

No potential conflict of interest was reported by the author(s).

Acknowledgments

The authors were thankful to Tamil Nadu Veterinary and Animal Sciences University Chennai, India, for providing the necessary infrastructure to carry out the research work.

References

- AL DAHOUK, S., H. TOMASO, E. PRENGER-BERNINGHOFF, W. D. SPLETTSTOESSER, H. C. SCHOLZ, H. NEUBAUER (2005): Identification of *Brucella* Species and Biotypes using Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP). *Crit. Rev. Microbiol.* 31, 191-196.
<https://doi.org/10.1080/10408410500304041>
- ALLEN, C. A., L. G. ADAMS, T. A. FICHT (1998): Transposon-derived *Brucella abortus* rough mutants are attenuated and exhibit reduced intracellular survival. *Infect. Immun.* 66, 1008-1016.
<https://doi.org/10.1128/IAI.66.3.1008-1016.1998>
- AURICH, S., J. SCHNEIDER, H. BRANGSCH, A. KOETS, F. MELZER, C. EWERS, E. PRENGER-BERNINGHOFF (2023): *Brucella suis* biovar 1 infection in a dog with orchitis in Germany. *Front. Vet. Sci.* 10, 1233118.
<https://doi.org/10.3389/fvets.2023.1233118>
- AZIZ, R. K., D. BARTELS, A. A. BEST, M. DEJONGH, T. DISZ, R. A. EDWARDS, K. FORMSMA, S. GERDES, E. M. GLASS, M. KUBAL, F. MEYER, G. J. OLSEN, R. OLSON, A. L. OSTERMAN, R. A. OVERBEEK, L. K. MCNEIL, D. PAARMANN, T. PACZIAN, B. PARRELLO, G. D. PUSCH, C. REICH, R. STEVENS, O. VASSIEVA, V. VONSTEIN, A. WILKE, O. ZAGNITKO (2008): The RAST Server: rapid annotations using subsystems technology. *BMC genomics.* 9, 75.
<https://doi.org/10.1186/1471-2164-9-75>
- BRICKER, B. J., S. M. HALLING (1994): Differentiation of *Brucella abortus* (biovars 1, 2, and 4), *Brucella melitensis*, *Brucella ovis*, and *Brucella suis* (biovar 1) by the polymerase chain reaction. *J. Clin. Microbiol.* 32, 2660-2666.
<https://doi.org/10.1128/jcm.32.11.2660-2666.1994>
- CORONAS-SERNA, J. M., A. LOUCHE, M. RODRÍGUEZ-ESCUADERO, M. ROUSSIN, P. R. C. IMBERT, I. RODRÍGUEZ-ESCUADERO, L. TERRADOT, M. MOLINA, J. P. GORVEL, V. J. CID, S. P. SALCEDO (2020): The TIR-domain containing effectors BtpA and BtpB from *Brucella abortus* impact NAD metabolism. *PLoS pathogens.* 16, e1007979.
<https://doi.org/10.1371/journal.ppat.1007979>
- DI SABATINO, D., G. GAROFOLO, A. DI PROVVIDO, K. ZILLI, G. FOSCHI, E. DI GIANNATALE, M. CIUFFETELLI, F. DE MASSIS (2017): *Brucella suis* biovar 2 multi locus sequence type ST16 in wild boars (*Sus scrofa*) from Abruzzi region, Italy. Introduction from Central-Eastern Europe? *Infect. Genet. Evol.* 55, 63-67.
<https://doi.org/10.1016/j.meegid.2017.08.031>
- FOSTER, J. T., S. M. BECKSTROM-STERNBERG, T. PEARSON, J. S. BECKSTROM-STERNBERG, P. S. CHAIN, F. F. ROBERTO, J. HNATH, T. BRETTIN, P. KEIM (2009): Whole-genome-based phylogeny and divergence of the genus *Brucella*. *J. Bacteriol.* 191, 2864-2870.
<https://doi.org/10.1128/JB.01581-08>
- GARDNER, S. N., T. SLEZAK, B. G. HALL (2015): kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 31, 2877-2878.
<https://doi.org/10.1093/bioinformatics/btv271>
- GARIN-BASTUJI, B., J. HARS (2001): Situation épidémiologique de la brucellose à *Brucella suis* biovar 2 en France. *Bull. Epidemiol.* 2, 3-4.
- GIRAULT, G., V. DJOKIC, F. PETOT-BOTTIN, L. PERROT, B. THIBAUT, H. SÉBASTIEN, A. F. VICENTE, C. PONSART, L. FREDDI (2023): Molecular Investigations of Two First *Brucella suis* Biovar 2 Infections Cases in French Dogs. *Pathogens.* 12, 792.
<https://doi.org/10.3390/pathogens12060792>

- GIRAULT, G., L. PERROT, V. MICK, C. PONSART (2022): High-Resolution Melting PCR as Rapid Genotyping Tool for *Brucella* Species. *Microorganisms* 10, 336.
<https://doi.org/10.3390/microorganisms10020336>
- GRE'GOIRE, F., B. MOUSSET, D. HANREZ (2012): A serological and bacteriological survey of brucellosis in wild boar (*Sus scrofa*) in Belgium. *BMC Vet. Res.* 8, 80.
<https://doi.org/10.1186/1746-6148-8-80>
- KARTHIK, K., S. ANBAZHAGAN, P. THOMAS, M. ANANDA CHITRA, T. M. A. SENTHILKUMAR, R. SRIDHAR, G. DHINAKAR RAJ (2021): Genome Sequencing and Comparative Genomics of Indian Isolates of *Brucella melitensis*. *Front. Microbiol.* 12, 698069.
<https://doi.org/10.3389/fmicb.2021.698069>
- LE FLÈCHE, P., I. JACQUES M. GRAYON, S. ALDAHOUK, P. BOUCHON, F. DENOEUDE, K. NÖCKLER, H. NEUBAUER (2006): Evaluation and selection of tandem repeat loci for a *Brucella* MLVA typing assay. *BMC Microbiol.* 6, 9.
<https://doi.org/10.1186/1471-2180-6-9>
- LETUNIC, I., P. BORK (2007): Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127-128.
<https://doi.org/10.1093/bioinformatics/btl529>
- LÓPEZ-GOÑI, D., C. M. GARCÍA-YOLDI, M. J. MARÍN, M. J., DE MIGUEL, E. BARQUERO-CALVO, C. GUZMÁN-VERRI, D. ALBERT, B. GARIN-BASTUJI (2011): New Bruce-ladder multiplex PCR assay for the biovar typing of *Brucella suis* and the discrimination of *Brucella suis* and *Brucella canis*. *Vet. Microbiol.* 154, 152-155.
<https://doi.org/10.1016/j.vetmic.2011.06.035>
- MESUREUR, J., S. RANALDI, V. MONNIN, V. GIRARD, S. AREND, M. WELKER, D. O'CALLAGHAN, J. P. LAVIGNE, A. KERIEL (2016): 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, 449-452.
<https://doi.org/10.1128/JCM.02730-15>
- OIE, World Organization for Animal Health (2013): Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. [http://www.oie.int/fileadmin- min/Home/eng/Health_standards/tahm/2.08.05_PORCINE_BRUC.pdf](http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.08.05_PORCINE_BRUC.pdf).
- PAGE, A. J., C. A. CUMMINS, M. HUNT, V. K. WONG, S. REUTER, M. T. HOLDEN, M. FOOKES, D. FALUSH, J. A. KEANE, J. PARKHILL (2015): Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691-3693.
<https://doi.org/10.1093/bioinformatics/btv421>
- PAPIĆ, B., D. KUŠAR, B. KRT, M. OCEPEK, J. AVBERŠEK (2023): Phylogeography of *Brucella suis* biovar 2 with focus on Slovenian wildlife. *Vet. Microbiol.* 282, 109751.
<https://doi.org/10.1016/j.vetmic.2023.109751>
- RASKO, D. A., M. J. ROISOVITZ, G. S. MYERS, E. F. MONGODIN, W. F. FRICKE, P. GAJER, J. CRABTREE, M. SEBAIHA, N. R. THOMSON, R. CHAUDHURI, I. R. HENDERSON, V. SPERANDIO, J. RAVEL (2008): The pan-genome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* 190, 6881-6893.
<https://doi.org/10.1128/JB.00619-08>
- SALCEDO, S. P., M. I. MARCHESINI, C. DEGOS, M. TERWAGNE, K. VON BARGEN, H. LEPIDI, C. K. HERRMANN, T. L. SANTOS LACERDA, P. R. IMBERT, P. PIERRE, L. ALEXOPOULOU, J. J. LETESSON, D. J. COMERCI, J. P. GORVEL (2013): BtpB, a novel *Brucella* TIR-containing effector protein with immune modulatory functions. *Front. Cell. Infect. Microbiol.* 8, 28.
<https://doi.org/10.3389/fcimb.2013.00028>
- SANKARASUBRAMANIAN, J., U. S. VISHNU, J. SRIDHAR, P. GUNASEKARAN, J. RAJENDHRAN (2015): Pan-Genome of *Brucella* Species. *Indian J. Microbiol.* 55, 88-101.
<https://doi.org/10.1007/s12088-014-0486-4>
- SEEMANN, T. (2019): snp-dists. In: GitHub repository. GitHub. <https://github.com/tseemann/snp-dists>
- SEEMANN, T. (2016a): MLST: scan contig files against PubMLST typing schemes. Github <https://github.com/tseemann/mlst>
- SEEMANN, T. (2016b): ABRicate: mass screening of contigs for antibiotic resistance genes. Github <https://github.com/tseemann/abricate>
- TAN, K. K., Y. C. TAN, L. Y. CHANG, K. W. LEE, S. S. NORE, W. Y. YEE, M. N. MAT ISA, F. L. JAFAR, C. C. HOH, S. ABUBAKAR (2015): Full genome SNP-based phylogenetic analysis reveals the origin and global spread of *Brucella melitensis*. *BMC Genomics* 16, 93.
<https://doi.org/10.1186/s12864-015-1294-x>
- TATUSOVA, T., M. DICUCCIO, A. BADRETDIN, V. CHETVERNIN, E. P. NAWROCKI, L. ZASLAVSKY, A. LOMSADZE, K. D. PRUITT, M. BORODOVSKY, J. OSTELL (2016): NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44, 6614-6624.
<https://doi.org/10.1093/nar/gkw569>
- TETTELIN, H., V. MASIGNANI, M. J. CIESLEWICZ, C. DONATI, D. MEDINI, N. L. WARD, S. V. ANGIUOLI, J. CRABTREE, A. L. JONES, A. S. DURKIN, R. T. DEBOY, T. M. DAVIDSEN, M. MORA, M. SCARSELLI, I. M. Y. ROS, J. D. PETERSON, C. R. HAUSER, J. P. SUNDARAM, W. C. NELSON, R. MADUPU, L. M. BRINKAC, R. J. DODSON, M. J. ROISOVITZ, S. A. SULLIVAN, S. C. DAUGHERTY, D. H. HAFT, J. SELENGUT, M. L. GWINN, L. ZHOU, N. ZAFAR, H. KHOURI, D. RADUNE, G. DIMITROV, K. WATKINS, K. J. O'CONNOR, S. SMITH, T. R. UTTERBACK, O. WHITE, C. E. RUBENS, G. GRANDI, L. C. MADOFF, D. L. KASPER, J. L. TELFORD, M. R. WESSELS, R.

- RAPPUOLI, C. M. FRASER (2005): Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. Proc. Natl. Acad. Sci. U S A 102, 13950-13955.
<https://doi.org/10.1073/pnas.0506758102>
- TONKIN-HILL, G., N. MACALASDAIR, C. RUIS, A. WEIMANN, G. HORESH, J. A. LEES, R. A. GLADSTONE, S. LO, C. BEAUDOIN, R. A. FLOTO, S. D. W. FROST, J. CORANDER, S. D. BENTLEY, J. PARKHILL (2020): Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biol. 21,180.
<https://doi.org/10.1186/s13059-020-02090-4>
- TRIFINOPOULOS, J., L. T. NGUYEN, A. VON HAESELER, B. Q. MINH (2016): W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232-W235.
<https://doi.org/10.1093/nar/gkw256>
- WANG, Y., Z. WANG, X. CHEN, H. ZHANG, F. GUO, K. ZHANG, H. FENG, W. GU, C. WU, L. MA, T. LI, C. CHEN, S. GAO (2016): The Complete Genome of *Brucella suis* 019 Provides Insights on Cross-Species Infection. Genes (Basel) 7, 7.
<https://doi.org/10.3390/genes7020007>
- WHATMORE, A. M., L. L. PERRETT, A. P. MACMILLAN (2007): Characterisation of the genetic diversity of *Brucella* by multilocus sequencing. BMC Microbiol. 7, 34.
<https://doi.org/10.1186/1471-2180-7-34>
- WICK, R. R., L. M. JUDD, C. L. GORRIE, K. E. HOLT (2017): Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoSComput. Biol. 13, e1005595.
<https://doi.org/10.1371/journal.pcbi.1005595>
- ZHAO, Y., X. JIA, J. YANG, Y. LING, Z. ZHANG, J. YU, J. WU, J. XIAO (2014): PanGP: a tool for quickly analyzing bacterial pan-genome profile. Bioinformatics 30, 1297-1299.
<https://doi.org/10.1093/bioinformatics/btu017>

Received: 30 May 2023

Accepted: 15 November 2023

Online publication: 10 May 2024

KARTHIK, K., S. ANBAZHAGAN, S. MANOHARAN: Usporedna genomska analiza u diferencijaciji različitih biovarova bakterije *Brucella suis*. Vet. arhiv 94, 285-296, 2024.

SAŽETAK

Soj *Brucella suis* TANUVAS_1 izoliran je u svinja s pobačajem te je sekvenciran cijeli genom tog soja. U istraživanju je provedena usporedna genomska analiza sojeva *B. suis* (n=33), uključujući novosekvencirani soj TANUVAS_1, kako bi se utvrdila točna diferencijacija biovarova. Da bi se ublažili nedostaci konvencionalnih tehnika identifikacije biovarova, u ovom je radu primijenjen cjelovit pristup genomu. Filogenetskom su analizom cijeloga genoma sojevi *B. suis* jasno diferencirani na osnovi biovarova. Potvrđeno je postojanje ukupno 43 gena virulencije, među kojima je gen *btpA* bio prisutan u samo 14 sojeva, dok je gen *wbkA* bio prisutan u 19 sojeva. Tipizacija sekvenciranjem na više lokusa (MLST) pokazala je pet različitih sekvencijskih tipova (14, 15, 16, 17 i 19), a najveći je broj sojeva pripadao ST-u 14. MLST analizom genoma ustanovljeno je 6 različitih ST-ova, a najveći je broj sojeva pripadao ST-u 50. Pangenomska analiza sojeva pokazala je da je ukupno 1948 gena (30,47%) pripadalo osnovnim genima, a 4445 gena (69,53%) bili su akcesorni geni, upućujući na to da genom *B. suis* ima otvorenu pangenomsku strukturu. Analiza jednonukleotidnih polimorfizama (SNP) otkrila je 17 SNP-ova specifičnih za biovar 1, 2142 SNP-a specifična za biovar 2, 673 SNP-a specifična za biovar 3, 714 SNP-ova specifičnih za biovar 4 i 2308 SNP-ova specifičnih za biovar 5. Rezultati su pokazali da se cgMLST može upotrijebiti za preciznu diferencijaciju sojeva na temelju biovarova. Također, SNP-ovi specifični za biovar mogu poslužiti za razvoj molekularnih testova kojima se biovarovi diferenciraju.

Cljučne riječi: *Brucella suis*; usporedna genomska analiza; geni specifični za biovar; SNP-ovi specifični za biovar; geni virulencije
