

Evolutionary history and current distribution of the West Mediterranean lineage of *Brucella melitensis* in Italy

Anna Janowicz¹, Fabrizio De Massis¹, Katuscia Zilli¹, Massimo Ancora¹, Manuela Tittarelli¹, Flavio Sacchini¹, Elisabetta Di Giannatale¹, Jason W. Sahl², Jeffrey T. Foster² and Giuliano Garofolo^{1,*}

Abstract

Ovine and caprine brucellosis, caused by *Brucella melitensis*, is one of the world's most widespread zoonoses and is a major cause of economic losses in domestic ruminant production. In Italy, the disease remains endemic in several southern provinces, despite an ongoing brucellosis eradication programme. In this study, we used whole-genome sequencing to detail the genetic diversity of circulating strains, and to examine the origins of the predominant sub-lineages of *B. melitensis* in Italy. We reconstructed a global phylogeny of *B. melitensis*, strengthened by 339 new whole-genome sequences, from Italian isolates collected from 2011 to 2018 as part of a national livestock surveillance programme. All Italian strains belonged to the West Mediterranean lineage, which further divided into two major clades that diverged roughly between the 5th and 7th centuries. We observed that Sicily serves as a brucellosis burden hotspot, giving rise to several distinct sub-lineages. More than 20 putative outbreak clusters of ovine and caprine brucellosis were identified, several of which persisted over the 8 year survey period despite an aggressive brucellosis eradication campaign. While the outbreaks in Central and Northern Italy were generally associated with introductions of single clones of *B. melitensis* and their subsequent dissemination within neighbouring territories, we observed weak geographical segregation of genotypes in the southern regions. Biovar determination, recommended in routine analysis of all *Brucella* strains by the World Organisation for Animal Health (OIE), could not discriminate among the four main global clades. This demonstrates a need for updating the guidelines used for monitoring *B. melitensis* transmission and spread, both at the national and international level, and to include whole-genome-based typing as the principal method for identification and tracing of brucellosis outbreaks.

DATA SUMMARY

Sequence reads for all 339 isolates were deposited in the Sequence Read Archive (SRA) under BioProject accession number PRJNA615379. A full list of SRA sample accession numbers and the associated metadata are provided in Table S1 (available with the online version of this article).

INTRODUCTION

Brucella melitensis is the causative agent of ovine and caprine brucellosis, and it is the most common cause of brucellosis in humans [1–3]. Small ruminants, including sheep and goats, are the natural hosts, but *B. melitensis* can also infect other animals, such as cattle (*Bos taurus*), water

buffalo (*Bubalus bubalis*) and camels (*Camelus* spp.) [4–6]. Brucellosis is among the most widespread zoonotic diseases worldwide, yet it is still considered a neglected zoonosis due to the limited resources devoted to its surveillance and control, particularly in economically developing countries [7–9]. Although infection with *B. melitensis* in humans is rarely fatal, it remains a major public-health concern, and in misdiagnosed and untreated patients it can progress to a chronic and debilitating disease with severe complications [10–12].

Ovine and caprine brucellosis likely first established in the Fertile Crescent, where wild goats and sheep were first domesticated [13–16]. From there, three main

Received 23 June 2020; Accepted 14 September 2020; Published 08 October 2020

Author affiliations: ¹National and OIE Reference Laboratory for Brucellosis, Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise "G. Caporale", via Campo Boario, 64100 Teramo, Italy; ²Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, AZ 86011, USA.

*Correspondence: Giuliano Garofolo, g.garofolo@izs.it

Keywords: *Brucella melitensis*; brucellosis; molecular epidemiology; WGS.

Abbreviations: cgMLST, core-genome multilocus sequence typing; EU, European Union; ML, maximum-likelihood; MLST, multilocus sequence typing; MST, minimum spanning tree; OIE, World Organisation for Animal Health; ST, sequence type; WGS, whole-genome sequencing.

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. One supplementary table is available with the online version of this article.

000446 © 2020 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

B. melitensis clades developed: an Americas group that includes strains from Spain, Portugal and much of North and South America, an East Mediterranean lineage that extends from the Middle East across to East Asia, and a West Mediterranean lineage found largely in Italy and the Maghreb [17–20]. An additional lineage, the African clade, diverged from the Americas group, and is found primarily in East Africa and the Arabian Peninsula [21]. Ovine and caprine brucellosis have now been eliminated from much of the Western world, but it is still present in domestic and wild animals in the Mediterranean Basin, Western Asia, and parts of Africa and South America, with re-emergence in the Middle East [22–24]. In the European Union (EU), the disease remains present in small ruminant flocks in several countries, including Croatia, Greece, Italy, Spain and Portugal [25].

Ewes and does infected with brucellosis suffer from reproductive failures, while male ruminants present with orchitis, epididymitis and infertility. Aborted fetuses, placentas and mucosal secretions, as well as milk, of chronically infected animals are the main transmission sources of the disease within flocks and to humans [26, 27]. Outbreaks of brucellosis in farms pose a risk to human health, and can lead to major economic losses for farmers and livestock breeders [28, 29]. Thus, in order to eliminate ovine and caprine brucellosis, Italy adopted a strict eradication programme in 1992, based on a test and slaughter policy [30, 31]. Since the introduction of the national plan, a slow but steady decrease in disease prevalence was observed [32]. In the 20 years from 1997 to 2017, brucellosis-positive farms declined from 3.9 to 1.0%, and several northern regions were declared ‘officially free from *B. melitensis*’ [30, 33]. Unfortunately, despite the surveillance and control measures introduced in the campaign, the disease still remains endemic in several southern provinces of Italy [19, 32–34]. Annually, the highest prevalence in sheep and goats is observed in Sicily, with 7604 animals in 208 flocks testing seropositive to *B. melitensis* in 2018 [35]. In the year 2016, an outbreak of human brucellosis was traced back to consumption of local soft cheese in Messina, during which 137 cases were reported; thus, highlighting the need for increased efforts to break transmission chains, especially in the most affected areas of the country [36].

In this work, we used whole-genome sequencing (WGS) to characterize *B. melitensis* strains circulating in Italy from 2011 to 2018 and to place sequenced genomes within a global phylogeny. Our approach used Bayesian temporal analysis to understand the evolutionary history and timing of the West Mediterranean lineage in Italy. Additionally, we used the computationally simplified method core-genome multilocus sequence typing (cgMLST) to determine the diversity and epidemiological relatedness of the current Italian *B. melitensis*. For surveillance purposes, data clustering using minimum spanning trees (MSTs) with the established cut-offs was performed to allow contact tracing of the outbreaks.

Impact Statement

Ovine and caprine brucellosis is a major zoonotic disease transmitted to humans most commonly via contaminated milk or dairy products. In our study, we sequenced and analysed 339 Italian strains of *Brucella melitensis*; therefore, increasing the publicly available genomes belonging to the West Mediterranean lineage, which to date has been under-represented. We established a possible period of diversification of *B. melitensis* clades in Italy more than a millennium ago, and described the recent epidemiology of ovine and caprine brucellosis in Italy over a period of 8 years. We demonstrated the use of whole-genome-sequencing methods to monitor the spread of clonal complexes at the national level, which revealed that the eradication campaign implemented in Italy has not been effective in eliminating some of the circulating clones, especially in the regions with the highest prevalence of brucellosis. Such hotspots of infection must, therefore, be considered as high-priority targets in order to eliminate the risk of brucellosis transmission to the other regions.

METHODS

B. melitensis isolate sequences

We sequenced 339 Italian *B. melitensis* isolates for this study; sequencing reads were deposited in the National Center for Biotechnology Information (NCBI) repository under BioProject accession number PRJNA615379. The dataset was supplemented by published sequences of 102 Italian strains [37] and 88 strains collected in Sweden [38]. An additional 149 genomes of *B. melitensis* with a reported isolation date were downloaded from the NCBI (accessed May 17 2019). The list of samples and the associated metadata are shown in Table S1.

Isolation of *B. melitensis*

Animal samples were collected after slaughtering livestock found positive for brucellosis in the context of the national eradication campaign from 2011 to 2018. Lymphatic glands, uterus or udder, and spleen samples were collected during necropsy. The isolates were cultured and identified by the *Istituto Zooprofilattico Sperimentale* laboratories using the standard protocol described in the World Organisation for Animal Health (OIE) Manual of Diagnostic Tests and Vaccines for Terrestrial Animals [4]. Human *B. melitensis* isolates were obtained from the blood samples of patients suspected of suffering from brucellosis. For DNA extraction, all of the *B. melitensis* isolates were sub-cultured in Brucella medium base and incubated in a 5–10% CO₂ atmosphere at 37°C for 48 h to assess the purity of cultures and the absence of dissociation. DNA extraction was performed using a Maxwell 16 tissue DNA purification kit (Promega), following the manufacturer’s instructions. The isolates were identified

as *B. melitensis* by PCR and biovar designation was obtained by standard biotyping as recommended by the OIE [4, 6].

WGS

Total genomic DNA from the 339 samples was sequenced using the Illumina NextSeq 500 platform. Briefly, following total DNA quantification (Qubit DNA HS assay; Thermo Fisher Scientific), the sequencing libraries were generated using a Nextera XT library preparation kit (Illumina), according to the manufacturer's instructions. The libraries were sequenced using NextSeq 500/550 mid output reagent cartridge v2 (300 cycles). Paired-end 150 bp reads were generated and their quality was assessed with FastQC [39]. Reads were trimmed with Trimmomatic version 0.36 and subsequently used to assemble genome scaffolds using SPAdes version 3.11.1 [40, 41].

Multilocus sequence typing (MLST) and cgMLST

Genome assemblies of Italian *B. melitensis* were genotyped using cgMLST. We used cgMLST because the method generates data that are relatively quick and simple to analyse, and the results can be readily standardized and reported. The cgMLST profiles were assigned using the *B. melitensis* task template with 2704 target core genes in Ridom SeqSphere+ software, version 4.1.1 (Ridom), as previously described [37]. Only genomes containing $\geq 98\%$ of good targets were accepted for the subsequent analyses. To be accepted as good targets, identified genes need to fulfil the software's default Target QC parameters, i.e. the same length as the reference gene ± 3 nucleotide triplets, no ambiguities and no frameshifts. A MST was constructed by pairwise comparison of cgMLST alleles. Based on previous data, the cut-off of six allele differences was applied to identify clusters of possibly related genomes [37]. Default parameters were used and the missing values were excluded in the calculation of distance between genotypes. The public genome dataset and the Italian strains sequenced in this study were genotyped with the *Brucella* nine locus MLST (MLST-9) scheme available at <https://pubmlst.org/brucella/> [42] and accessed through Ridom SeqSphere+.

SNP detection

SNPs were called using In Silico Genotyper (iSG) (version 0.16.10-3), with BWA-MEM (version 0.7.12-r1039) as the aligner and GATK (version 3.9) as the SNP caller using *B. melitensis* strain 16M (GenBank accession numbers NC_003317.1 and NC_003318.1) as the reference [43-45]. Default filters were applied to remove SNPs from duplicated regions, minimum quality was set to Phred 30 and the minimum allele frequency was set to 90%. The alignments of concatenated, clean and unique variants were used in further analyses.

Phylogenetic analysis

A maximum-likelihood (ML) tree was built from the sequence ($n=499$) alignment of 14531 concatenated SNPs called by iSG. RAxML-NG version 0.9.0 was used to perform the analysis with the generalized time-reversible model with gamma distribution and base frequencies optimized by ML

[46]. The number of invariant sites was accounted for in the model and Stamatakis correction was used to correct for ascertainment bias. Branch support for the phylogeny was assessed using the transfer bootstrap expectation method with 650 bootstrap replicates (number of sufficient bootstrap replicates was estimated by a majority rules extended (MRE) based bootstrapping test performed in RAxML-NG. The best tree was mid-point rooted and visualized using version 4.4.2 of the Interactive Tree of Life (iTOL) online tool [47]. Population structure was assessed using BAPS 6.0 [in the module hierarchical BAPS (hierBAPS)] [48].

Temporal analysis

Divergence times for *B. melitensis* strains were estimated using the Bayesian Markov chain Monte Carlo method implemented in the program BEAST version 2.6.1 [49]. We first subsampled our dataset to reduce redundancy of the over-represented clades and, therefore, minimize sampling bias. The dataset of 259 samples was obtained by pruning the ML phylogenetic tree to 0.90 relative length with Treemmer [50]. The analysis was performed using an alignment of concatenated SNPs obtained using the iSG tool. Two molecular clock models, strict and relaxed, were tested in combination with coalescent constant, coalescent exponential and Bayesian skyline demographic models using stepping stone/path sampling analysis in BEAST 2 (BEAST Path Sampler in Model Selection package, version 1.5.2). The BEAST 2 package bModelTest version 1.2.1 was used to select the optimal base substitution model in all of the analyses [51].

The strict clock model with the coalescent exponential demographic model best fit for our dataset, as previously shown for *B. melitensis* [16]. Three independent Markov chain Monte Carlo analyses were performed, each with a chain length of 200 million generations. Invariant site values were provided manually in the xml Beauti file. The first 40 million chains were discarded as a burn-in and posterior distributions were sampled every 20000 generations. Log files were inspected using Tracer version 1.7.1 to ensure convergence and effective sample size above 200. The log and tree files of the three runs were combined using LogCombiner version 2.6.1 (Prerelease). Maximum clade credibility trees were generated using TreeAnnotator version 2.6.1 (Prerelease) and visualized using FigTree version 1.4.1.

RESULTS

Global phylogeny of *B. melitensis*

Until recently, only a few Italian *B. melitensis* genomes were publicly available, so the structure of this lineage remained poorly known. Our sequencing of 339 strains enabled a detailed reconstruction of the placement of Italian strains within a global phylogeny (Fig. 1). We identified 15041 SNPs and used them to build a phylogenetic tree using ML. The characteristic topology of the *B. melitensis* phylogeny was observed, with Italian strains clustered within the West

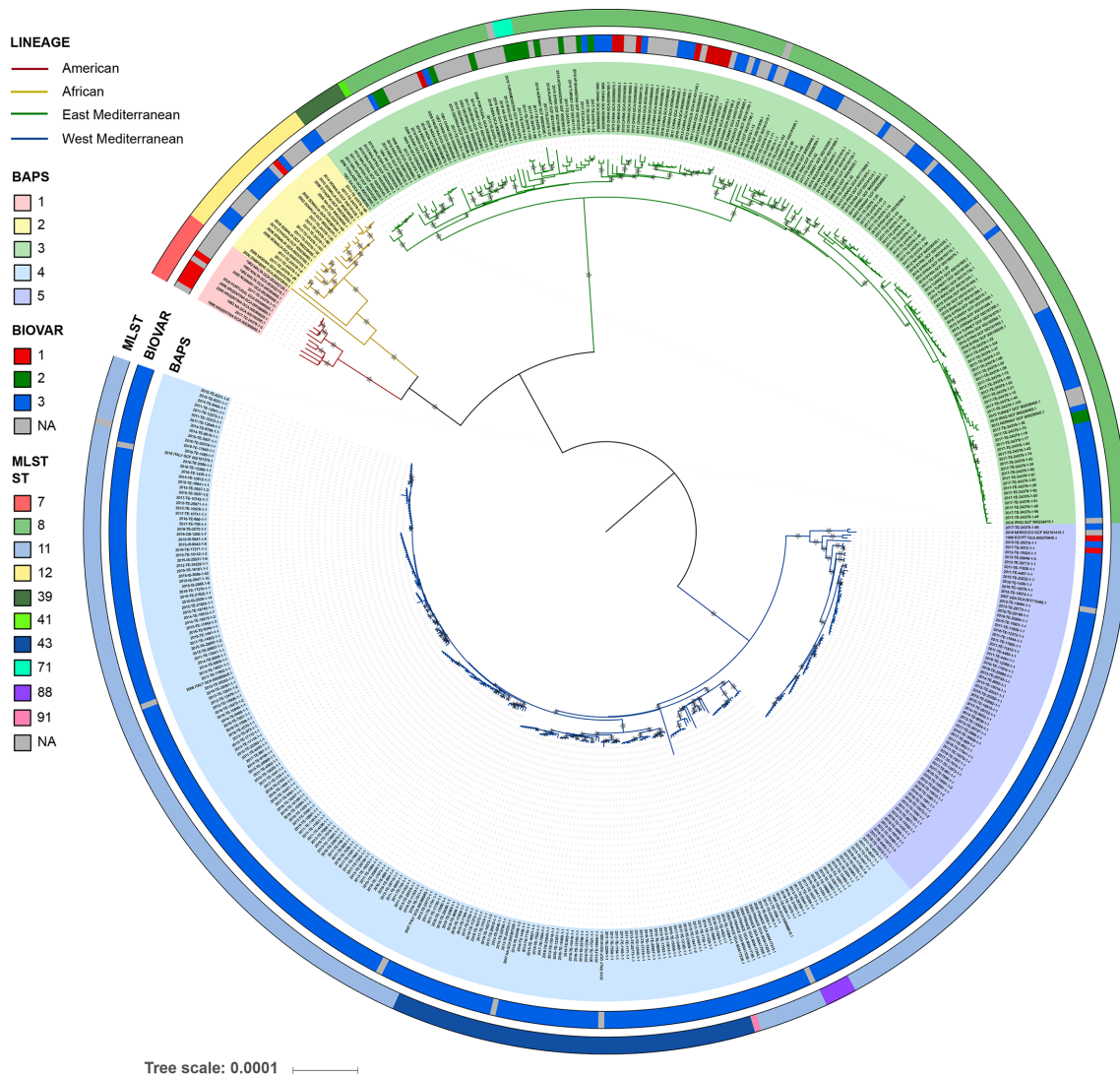


Fig. 1. Global phylogeny of *B. melitensis*. The ML tree was generated from an alignment of 15041 concatenated SNPs from 499 *B. melitensis* genomes. The branch colours correspond to the four main lineages and the strain ID labels are shaded according to the assigned BAPS population. The two outside rings correspond to biovar and MLST genotype. The tree is mid-point rooted, and branches with bootstrap values between 0.99 and 1 are marked with a star. Bar, mean number of nucleotide substitutions per site.

Mediterranean lineage, and three other phylogeographical lineages evident (Americas, African and East Mediterranean). We discovered 1004 SNPs unique to the West Mediterranean clade. Two main branches were identified within this lineage, one containing strains from Egypt, Italy and Morocco, and the other with strains from France, Italy and the USA.

Based on a secondary level of clustering of the whole-genome SNPs, a BAPS analysis split the tree into five lineages and confirmed the presence of two diverse groups within the West Mediterranean lineage (BAPS 4 and BAPS 5) (Fig. 2). The larger clade, BAPS 4, further divided into two subpopulations, one containing only Italian strains, and the other including strains from Italy and France. All Italian *B. melitensis* isolates belonged to biovar 3, which was the

most common biovar in our dataset, and was also assigned to a majority of strains from Africa and to a large subpopulation within the East Mediterranean lineage. Unfortunately, the majority of publicly available genomes lack biotyping results, so the relationship of biovar to genetic group could not be determined.

MLST analysis revealed four sequence types (STs) of *B. melitensis* present in Italy. The majority of the strains belonged to ST-11 and ST-43. ST-88 was generally confined to animals from the Apulia (Puglia) region and, interestingly, ST-91 was identified in only one isolate, which had been collected from a water buffalo in Campania. The latter isolate did not appear to belong to any sub-cluster, but formed a single long branch within the BAPS 4 population and had 270 unique SNPs.

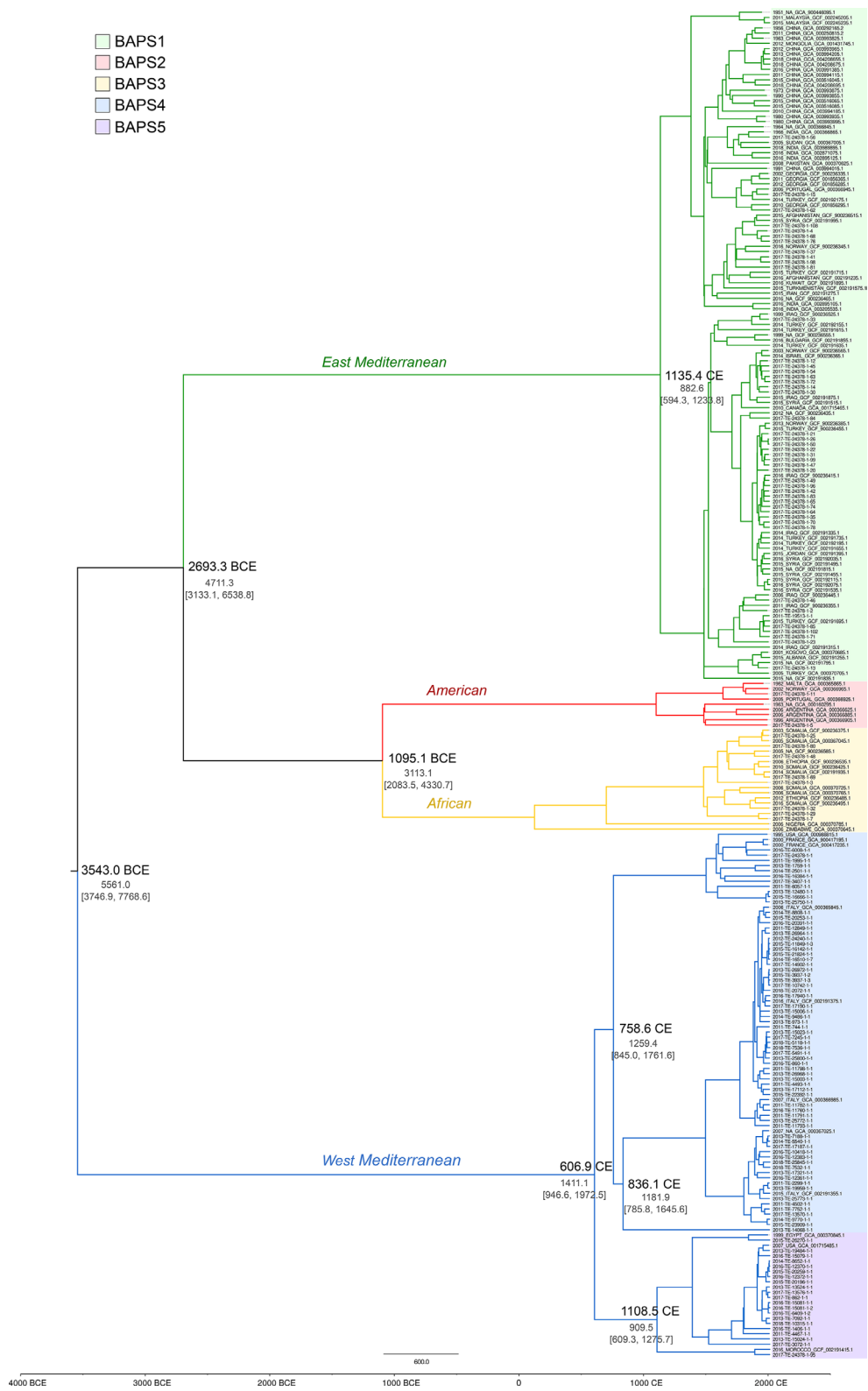


Fig. 2. Time-structured phylogeny of a subset of 259 isolates of *B. melitensis*. The maximum clade credibility tree was built using concatenated SNP sequences. The mean node ages are shown for the major nodes together with the node heights and corresponding 95% highest posterior density (HPD) intervals (in grey). The branches are coloured according to the lineage and the BAPS populations are highlighted with different colours as shown in the legend.

Origins and divergence of the West Mediterranean lineage of *B. melitensis*

According to the results of our temporal analysis, the West Mediterranean lineage was the first to diverge from the *B. melitensis* population, dividing into two branches in ~3500 BCE at the dawn of the Neolithic period, and after the first introductions of domesticated sheep and goats occurred in the Central Mediterranean region. The lineage split into two, which corresponded to BAPS 4 and BAPS 5 subclades, that shared a most recent common ancestor (MRCA) around 1400 years ago (95% confidence interval: 946.6–1972.5). Another major division appears to have occurred less than 200 years later, which resulted in division of BAPS 4 into two subpopulations. One of these was entirely composed of Italian isolates, while the other was shared with French strains.

Emergence of a distinct branch corresponding to the *B. melitensis* isolate from a water buffalo (2013-TE-14068-1-1) occurred around the same time as the main split of BAPS 4. The temporal evolution of the genetic clade corresponding to BAPS 5 involved two divergence events, the first occurring in approximately 1100 CE and the second roughly 200 years later. Both involved diversification of Italian strains from strains found in Africa.

Genomic epidemiology of ovine and caprine brucellosis in Italy

The majority of *B. melitensis* strains analysed in our study were isolated from animals in Sicily ($n=240$) and Calabria ($n=87$). In particular, the territories surrounding the Strait of Messina, which joins the two regions, had a particularly high burden of brucellosis (Table S1). Epidemiological investigation using cgMLST allowed us to estimate genetic relationships among the isolates within our dataset. The strains in the two BAPS populations present in Italy differed by more than 467 allelic differences. The larger clade, BAPS 4, contained isolates collected in 13 regions of Italy (Fig. 3), while BAPS 5 strains were found only in Sicily and Calabria (Fig. 4). Two subclades within the BAPS 4 population were separated by a distance of 367 allelic differences. The only strain assigned to ST-91 isolated from buffalo in 2014 differed by 372 alleles from the closest neighbouring strain, again highlighting its uniqueness.

We identified 21 cgMLST complexes containing a minimum of five strains and located within a maximum distance of six diverse core genes to the nearest neighbour (Table 1). Isolates assigned to these clusters were distributed across Italy, while the 'sporadic cases', i.e. not belonging to any complex, were found mainly in Sicily and Calabria. While these two regions, and Sicily in particular, are a hotspot of brucellosis, a clear geographical separation of the clonal clusters was observed in central and northern regions (Fig. 5). Epidemiological metadata suggests that some of the complexes might have originated in Sicily and then, likely due to the livestock trade, established in other parts of the country, resulting in local outbreaks of brucellosis. For example, we found several isolates from C2 that were collected in 2011 and 2012 in

Sicily and subsequently, in 2015 and 2016, the same clone was identified in an outbreak in Lazio and Molise. A similar scenario was observed for C4, present in Sicily in 2013 and 2014, which in later years also emerged in Lazio, Apulia and Campania. Conversely, few complexes, such as C8 and C20, contained only strains from mainland Italy.

The two largest clusters, C1 and C2, each containing approximately 10% of the sequenced isolates, persisted over the entire studied period. While the prevalence of the C2 clone dramatically decreased after detection of the outbreak and implementation of control measures in 2015 and 2016, the number of brucellosis cases caused by the C1 clone has been increasing (Fig. 5a). C1 isolates were traced to 31 farms found in several locations in the east of Sicily and in the south of Calabria, and were obtained from sheep, goats and cattle. Interestingly, the majority of the clonal complexes contained strains that were isolated from more than one animal species, and three complexes also contained *B. melitensis* genomes obtained from human samples.

In several of the largest complexes, we found that the difference between pairs of the most genetically distant isolates exceeded the cluster cut-off of six core genes. This observation was true even for the strains isolated within the same year, suggesting that they likely were not directly related, but might have shared a recent common ancestor.

DISCUSSION

Brucellosis remains one of the most widespread zoonoses in the world despite implementation of surveillance, control and eradication campaigns at the national and international scale. In the EU, the disease is endemic in sheep and goats in some of the southern member states, which annually record the highest number of human brucellosis cases [25]. Disease incidence in humans is further increased by infections acquired outside of Europe, mainly from the Middle East and Africa [38, 52, 53]. Between 2011 and 2018, the European Centre for Disease Prevention and Control (ECDC) reported 3645 human cases of brucellosis in the EU, 30% of which were diagnosed in Italy (dataset provided by the ECDC based on data provided by the World Health Organization and the Ministries of Health from the affected countries) [54]. In our study, we used WGS tools to analyse a large set of isolates collected during the same time period and provide a comprehensive picture of the epidemiology, genetic diversity and evolutionary history of *B. melitensis* in Italy.

To date, in the studies that used WGS to characterize the global population of *B. melitensis*, the West Mediterranean lineage was under-represented due to a lack of sequenced isolates [16, 53, 55]. In our work, we included recently published WGS data from Italy, as well as an additional 339 genomes; thus, expanding the phylogeny and providing a comprehensive analysis of the West Mediterranean lineage. The SNP-based phylogenetic analysis and subsequent BAPS group assignments divided this lineage into two diverse clades. Isolates from both populations were found

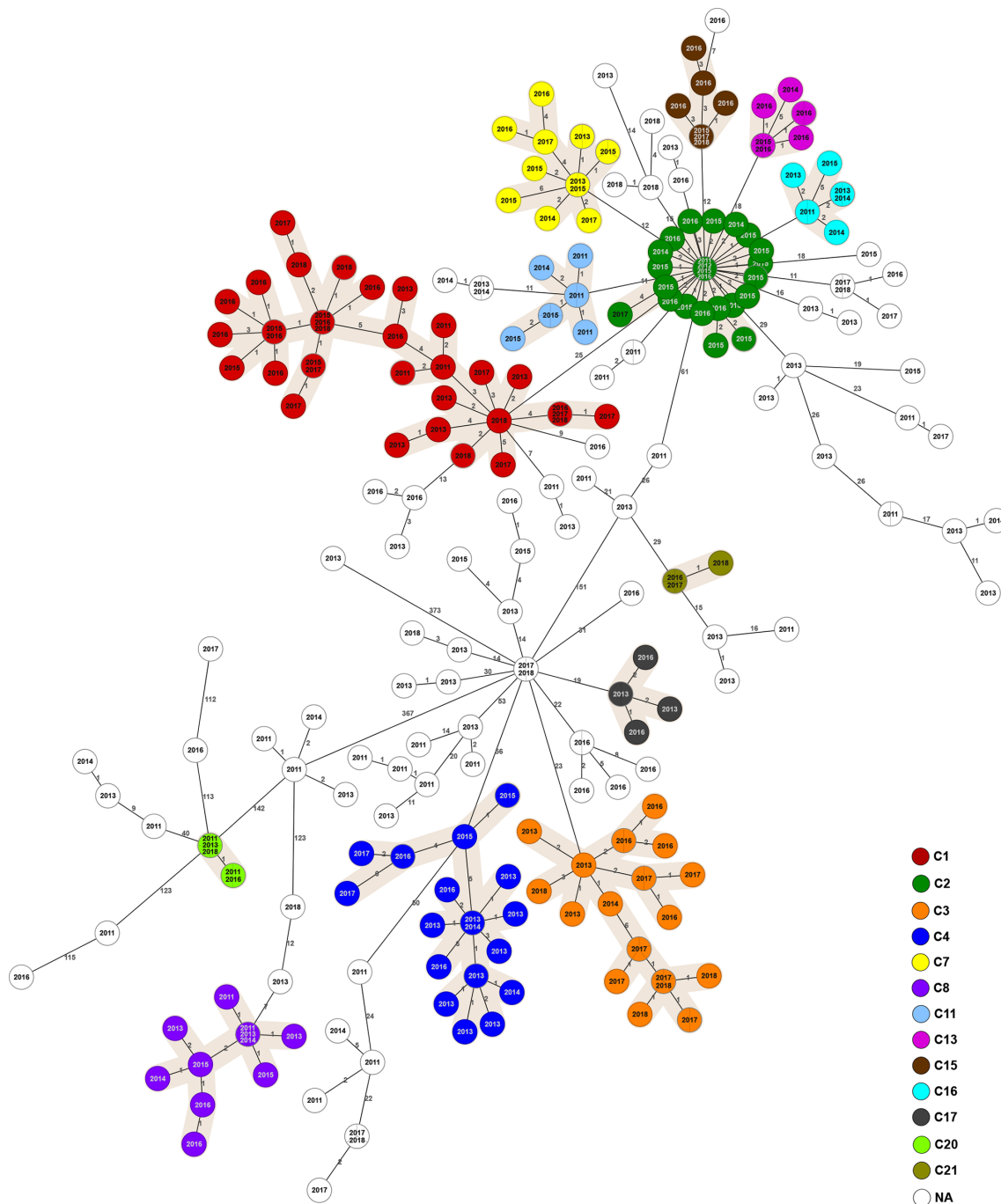


Fig. 3. MST generated for 317 Italian isolates of *B. melitensis* in the BAPS 4 population using the gene-by-gene approach. The MST was calculated by pairwise comparison of 2704 target genes with missing values ignored. Node labels correspond to strain isolation year and the branches to the number of discriminating loci. Complexes of genotypes within a distance of six alleles and containing a minimum of five isolates are numbered and depicted with different colours. Two main phylogenetic clusters are highlighted.

to circulate in Sicily and Calabria, but only BAPS 4 strains were responsible for outbreaks in the central and northern regions. The factors behind geographical restriction of BAPS 5 and more widespread distribution of the other clade are not clear; however, it is possible that the BAPS 4 isolates

were simply more numerous and, therefore, more likely to be transferred to other regions by trade.

Studies on the geographical partitioning of goat species revealed the presence of genetic groups related to the

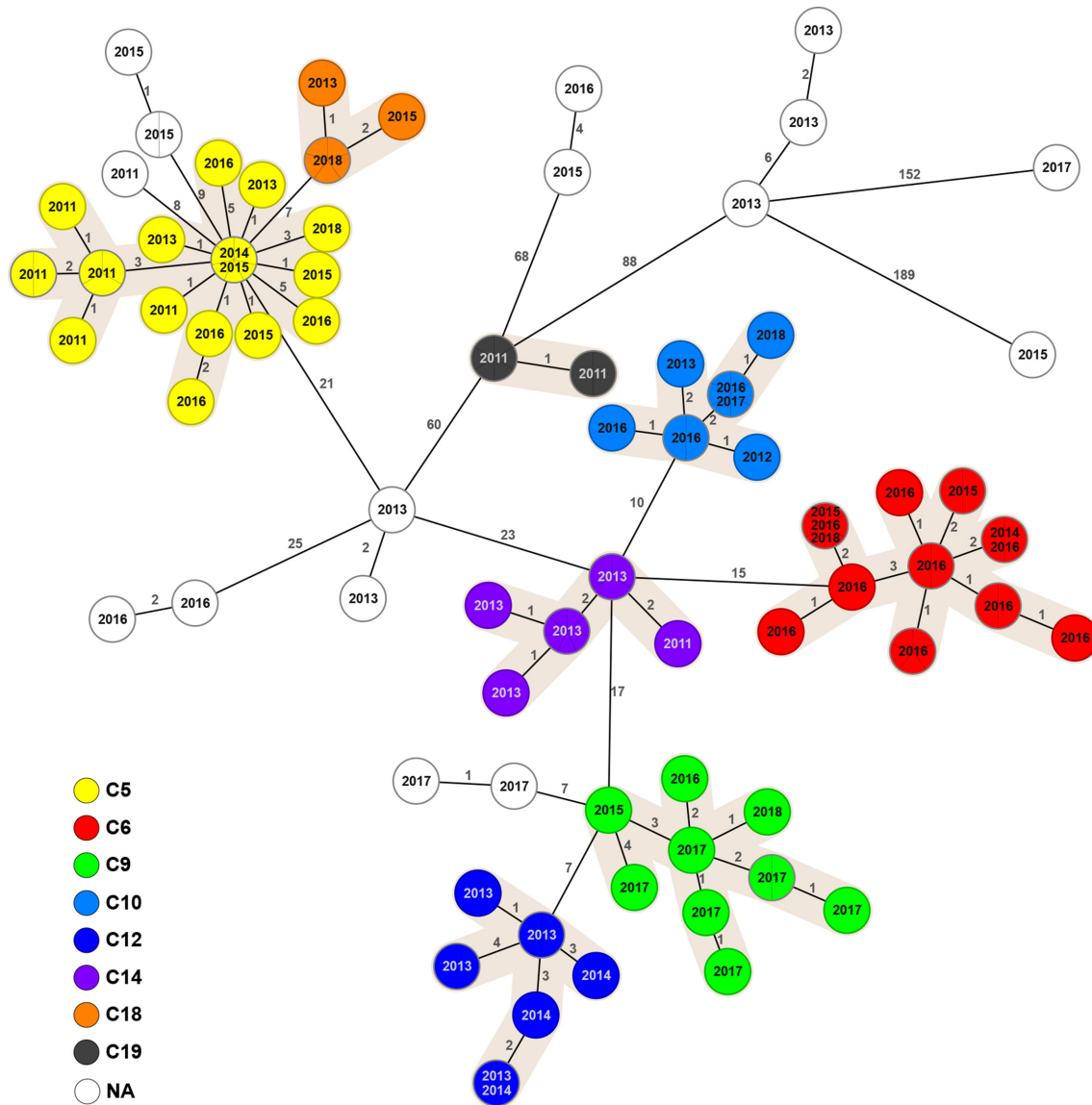


Fig. 4. MST generated for 123 Italian isolates of *B. melitensis* in the BAPS 5 population using the gene-by-gene approach. The MST was calculated by pairwise comparison of 2704 target genes with missing values ignored. Node labels correspond to strain isolation year and the branches to the number of discriminating loci. Complexes of genotypes within a distance of six alleles and containing a minimum of five isolates are numbered and depicted with different colours. Two main phylogenetic clusters are highlighted.

following geographical areas: the East, Central and West Mediterranean, and Central/Northern Europe [56, 57]. The genetic distances suggested that goats were first domesticated around 10500 years ago in the Fertile Crescent, from where they spread to Europe along with Neolithic human populations [58]. Similar observations have been made for sheep [59]. Our phylogenetic analysis of *B. melitensis* follows similar phylogeographical patterns. According to our temporal analysis, the West Mediterranean lineage was the first to diverge from the original East Mediterranean lineage more than 5500 years ago, which is similar in timing to previous estimates [16]. Our findings for *B. melitensis* population divergence coincide with the timing

of the Neolithic expansion, suggesting that the bacterium likely arrived in the central Mediterranean with infected sheep or goat populations. Then, the West Mediterranean lineage evolved in the region within its own species of ruminants. Although there is no archaeological evidence for the exact date of introduction of *B. melitensis* to Italy, we suspect that it happened 5000 years ago. Osteological evidence supports likely brucellosis in people from the archaeological sites of Pompei and Herculaneum in 79 BCE [60]. Furthermore, Kay and colleagues demonstrated the presence of the West Mediterranean lineage in Italy (Sardinia) using a metagenomics approach on a skeleton dated from the 1300s [61].

Table 1. *B. melitensis* cgMLST complexes present in Italy between 2011 and 2018

Complex	No. of isolates	No. of genotypes	No. of farms	2011/2012	2013/2014	2015/2016	2017/2018	Host	Region	Maximum distance cgMLST
1	44	22	31	x	x	x	x	Cattle, goat, sheep	Calabria, Sicily	17
2	46	21	24	x	x	x	x	Cattle, goat, human, sheep	Abruzzi, Lazio, Molise, Sardinia, Sicily	8
3	28	17	13			x	x	Cattle, goat, sheep	Emilia Romagna, Sicily, Tuscany	12
4	25	17	19		x	x		Cattle, goat, sheep	Campania, Lazio, Apulia, Sicily	19
5	24	15	14	x	x	x	x	Cattle, goat, sheep	Sicily	10
6	23	10	7		x	x	x	Goat, sheep	Calabria	6
7	14	10	9		x	x	x	Cattle, goat, sheep	Calabria, Sicily	11
8	14	9	9	x	x	x		Cattle, goat, human, sheep	Campania	5
9	10	9	7			x	x	Cattle, goat, sheep	Sicily	8
10	12	6	7	x	x	x	x	Goat, sheep	Calabria, Sicily	5
11	10	6	5	x	x	x		Goat, sheep	Calabria	5
12	9	6	7		x			Sheep	Sicily	5
13	17	5	4		x	x		Cattle, goat, sheep	Calabria	6
14	13	5	10	x	x			Goat, sheep	Sicily	5
15	8	5	7			x	x	Cattle, goat, sheep	Sicily	8
16	7	5	7	x	x	x		Goat, sheep	Sicily	7
17	5	4	2		x	x		Goat, sheep	Sicily	4
18	7	3	7		x	x	x	Cattle, goat, sheep	Sicily	3
19	8	2	7	x				Sheep	Sicily	1
20	6	2	4	x	x	x	x	Human, sheep	Abruzzi, Apulia	1
21	5	2	4			x	x	Cattle, goat, sheep	Sicily	1

In our study, we observed two major subsequent divisions of the West Mediterranean lineage that occurred around the year 600 CE, the first splitting it into two subclades corresponding to BAPS 4 and BAPS 5, and the second that further divided BAPS 4. Since both populations are now prevalent in Sicily,

at least two separate introduction events (or introduction and re-introduction) appear to have occurred in this region. In the BAPS 4 population, we observed two subpopulations. One of these was entirely composed of Italian isolates, while the other showed Italian isolates from the mainland along

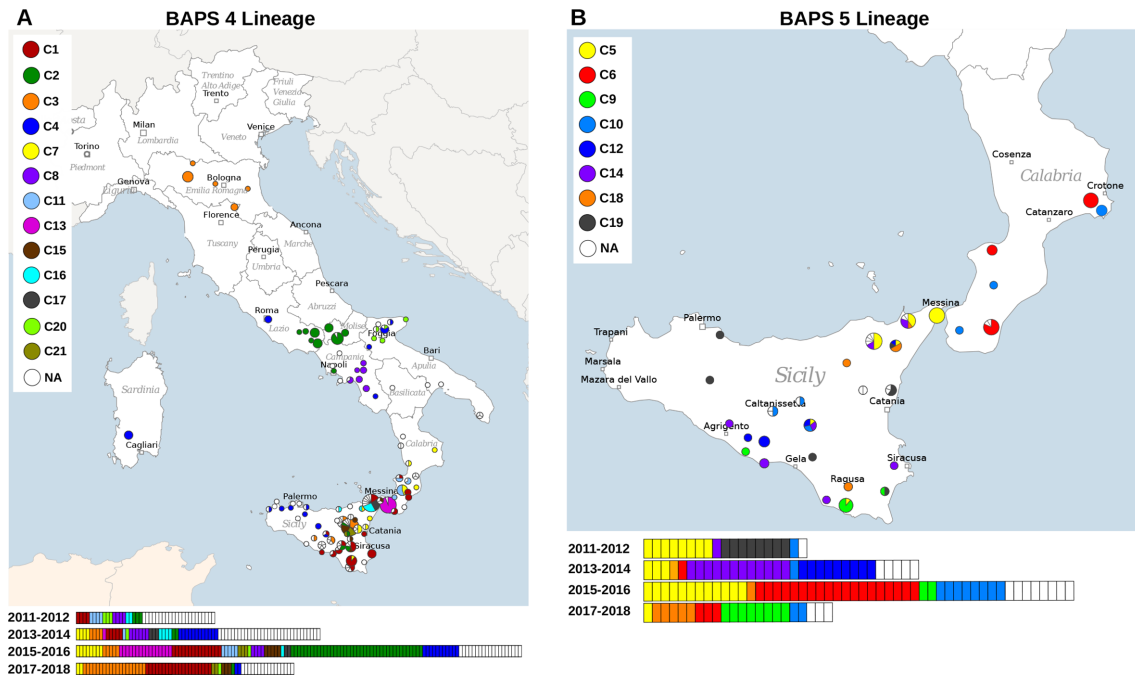


Fig. 5. Geographical distribution of cgMLST complexes of *B. melitensis* in Italy. A set of 440 isolates was typed using cgMLST and the clusters of at least five genotypes placed within six gene distance from the nearest neighbour were grouped into complexes. The isolates were assigned to BAPS populations using hierBAPS. The BAPS 4 strains are shown in (a) and the BAPS 5 strains are shown in (b). The isolates are coloured according to the assigned complex, and the numbers of individual samples coloured according to the assigned complex and isolated in specified time period are shown under the maps.

with French strains. The branches of the latter were long, suggesting an extended time for evolution within the cluster, partially explained by the geographical distribution of the sampled animals. Interestingly, we found that the Italian *B. melitensis* strain most closely related to the strain from France had been isolated from an alpine ibex from the Gran Paradiso National Park. The close localization of this park to the Bargy area in France suggests that this genotype is probably well established in France; however, further studies are required to understand the genetic diversity of *B. melitensis* found in this region. Within the BAPS 5 population, we identified two branches corresponding to strains from Egypt, Morocco and Eritrea that diverged from the Italian strains approximately 800–1000 years ago. The history of brucellosis spread in the Mediterranean basin could be explained by the trade connections between the southern Italian regions, Sicily in particular, and the Arabs and Berbers from Africa. While the West Mediterranean lineage may have been introduced to Italy in the Neolithic period, the spread of this lineage in Italy was probably associated with human movements. A strong trade connection between Italy and North Africa lasted from the time of the Roman Empire to the Middle Ages, and continued in the last century, in the period of modern colonialism. Commercial activities between both regions could, therefore, explain the history of brucellosis introductions between Africa and the Northern Mediterranean region [62]. At present, a limited number of strains from Africa and the lack of *B. melitensis* genomes from other European

countries within the Mediterranean region prevents us from deciphering the full history of ovine and caprine brucellosis in this region.

Core-genome analysis of the Italian *B. melitensis* population identified 21 complexes of closely related isolates, several of which had been circulating in Italy for at least 8 years. Our data demonstrate that the C1 cluster has been steadily expanding, despite an aggressive eradication campaign. Moreover, unlike in the highly clonal C2, the genetic distances between individual isolates in C1 could be as high as 17 core genes. Similar genetic distances were also found for another large complex, C4, which was detected between 2013 and 2017 in four different regions. These data would suggest that there are still gaps in the brucellosis detection system or in the process of elimination of the infected animals, leading to some strains remaining undetected for many years. These strains then persist in the animal populations and transmit to new farms when different flocks are grazing in the shared pastures or through the trading activities; thus, hindering the elimination of *B. melitensis* from Italy.

The presence of a long branch within the BAPS 4 population, corresponding to a single strain isolated from a water buffalo, raises questions about the origins of the animal or the bacterial strain. It is possible that this isolate was derived from a genetic type that has now been largely eliminated due to the previous brucellosis eradication efforts. In fact, water buffalo in Italy are more commonly infected with *B. abortus*

and the numbers of buffalo positive with *B. melitensis* have been steadily decreasing in the recent years [6, 32]. Analysis of genome sequences of *B. melitensis* strains circulating in Italy in the more distant past would be necessary in order to provide more information on the dynamics of emergence and extinction of specific lineages over time.

While using WGS in epidemiology and traceback of brucellosis cases has become a common practice, in particular in larger laboratories and reference centres, other methods such as biotyping are still recommended by the OIE and are, therefore, routinely performed. Here, the biotyping analysis showed that the Italian isolates of *B. melitensis* belong to biovar 3, but this biovar was also found in strains from the African and East Mediterranean lineages and, therefore, did not provide any additional epidemiological value for brucellosis investigations in Italy. MLST showed higher genotyping resolution than biotyping, as MLST-9 STs 11, 43, 88 and 91 were unique to the West Mediterranean clade and, therefore, could give an indication of the lineage. However, this typing method was not sufficient to distinguish between individual subclades.

WGS analysis, including cgMLST and whole-genome SNP typing, has superior genotyping resolution and is particularly well suited to large-scale epidemiological studies and disease surveillance monitoring when sufficient funding is available. However, it is important to note that in the case of insufficient sampling, or when few quality sequences are available, genetic linkage analyses might be misleading and genetically related sequences may not pass the cluster inclusion threshold. In our study, we included 441 Italian strains of *B. melitensis*; however, the actual number of positive animals between years 2011 and 2018 was several times higher [34]. Thus, it is possible that some connections between isolates or clusters were not sampled.

Ovine and caprine brucellosis has been a recognized problem in Italy for many decades. Despite the successful elimination of the disease in several northern regions of the country, the disease remains highly prevalent in Sicily and Calabria, affecting livestock production markets, limiting the live-animal trade and posing a lasting threat to human health. Moreover, the persistence of brucellosis in southern regions presents a continuous risk of spillover to the 'officially *B. melitensis*-free territories'. While the geography of affected regions and the farming practices remain major factors in the propagation of the disease, the expansion of some of the genetic *B. melitensis* complexes in recent years suggests the current surveillance and eradication programme is not effective enough for the complete elimination of the ovine and caprine brucellosis. An in-depth study of the factors that are involved in disease transmission and limit the effectiveness of control campaigns would be vital for development of a successful eradication plan in *B. melitensis* hotspots in Italy.

Funding information

This work has received funding from the Italian Ministry of Health under grant agreement codes IZS AM 02/17 RC and IZS AM 02/19 RC.

Acknowledgements

We thank the Italian Istituti Zooprofilattici Sperimentali network for their technical assistance with the *Brucella* diagnostics.

Author contributions

Conceptualization: G. G. Data curation: A. J. Formal Analysis: A. J. Funding acquisition: G. G., E. D. G., F. S., M. T. Investigation: K. Z., M. A. Methodology: A. J., G. G. Project administration: G. G. Resources: K. Z. Supervision: G. G. Visualization: A. J. Writing – original draft: A. J. Writing – review and editing: A. J., J. T. F., J. W. S., G. G., F. D. M.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Ethical statement

The National and OIE Reference Laboratory for Brucellosis at the Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise (IZSAM) is a public laboratory authorized by the Italian Ministry of Health to perform systematic brucellosis diagnosis. The *Brucella* isolates used in this retrospective study were stored in the strain collection at the IZSAM. All information and any other data that could potentially lead to the identification of individual patients or farm owners were anonymized. Therefore, we did not require any consent.

References

- Spink WW, Hall JW, Finstad J, Mallet E. Immunization with viable *Brucella* organisms. Results of a safety test in humans. *Bull World Health Organ* 1962;26:409–419.
- Whatmore AM. Current understanding of the genetic diversity of *Brucella*, an expanding genus of zoonotic pathogens. *Infect Genet Evol* 2009;9:1168–1184.
- Corbel MJ. Brucellosis: an overview. *Emerg Infect Dis* 1997;3:213–221.
- World Organisation for Animal Health. Brucellosis (*Brucella abortus*, *B. melitensis* and *B. suis*) (infection with *B. abortus*, *B. melitensis* and *B. suis*). *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. Paris: World Organisation for Animal Health; 2016.
- Refai M. Incidence and control of brucellosis in the Near East region. *Vet Microbiol* 2002;90:81–110.
- Di Giannatale E, De Massis F, Ancora M, Zilli K, Alessiani A. Typing of *Brucella* field strains isolated from livestock populations in Italy between 2001 and 2006. *Vet Ital* 2008;44:383–388.
- Godfroid J, Al Dahouk S, Pappas G, Roth F, Matope G et al. A "one health" surveillance and control of brucellosis in developing countries: moving away from improvisation. *Comp Immunol Microbiol Infect Dis* 2013;36:241–248.
- McDermott J, Grace D, Zinsstag J. Economics of brucellosis impact and control in low-income countries. *Rev Sci Tech* 2013;32:249–261.
- Zhang N, Huang D, Wu W, Liu J, Liang F et al. Animal brucellosis control or eradication programs worldwide: a systematic review of experiences and lessons learned. *Prev Vet Med* 2018;160:105–115.
- Franco MP, Mulder M, Gilman RH, Smits HL. Human brucellosis. *Lancet Infect Dis* 2007;7:775–786.
- Dean AS, Crump L, Greter H, Schelling E, Zinsstag J. Global burden of human brucellosis: a systematic review of disease frequency. *PLoS Negl Trop Dis* 2012;6:e1865.
- Zheng R, Xie S, Lu X, Sun L, Zhou Y et al. A systematic review and meta-analysis of epidemiology and clinical manifestations of human brucellosis in China. *Biomed Res Int* 2018;2018:5712920.
- Lv F-H, Peng W-F, Yang J, Zhao Y-X, Li W-R et al. Mitogenomic meta-analysis identifies two phases of migration in the history of eastern Eurasian sheep. *Mol Biol Evol* 2015;32:2515–2533.
- Pereira F, Amorim A. Origin and spread of goat pastoralism. *Encyclopedia of Life Sciences*. Chichester: Wiley; 2010.

15. **Moreno E.** Retrospective and prospective perspectives on zoonotic brucellosis. *Front Microbiol* 2014;5:213.
16. **Pisarenko SV, Kovalev DA, Volynkina AS, Ponomarenko DG, Rusanova DV et al.** Global evolution and phylogeography of *Brucella melitensis* strains. *BMC Genomics* 2018;19:353.
17. **Al Dahouk S, Flèche PL, Nöckler K, Jacques I, Grayon M et al.** Evaluation of *Brucella* MLVA typing for human brucellosis. *J Microbiol Methods* 2007;69:137–145.
18. **Whatmore AM, Koylass MS, Muchowski J, Edwards-Smallbone J, Gopaul KK et al.** Extended multilocus sequence analysis to describe the global population structure of the genus *Brucella*: phylogeography and relationship to biovars. *Front Microbiol* 2016;7:2049.
19. **Garofolo G, Ancora M, Di Giannatale E.** MLVA-16 loci panel on *Brucella* spp. using multiplex PCR and multicolor capillary electrophoresis. *J Microbiol Methods* 2013;92:103–107.
20. **Lounes N, Cherfa MA, Le Carrou G, Bouyoucef A, Jay M et al.** Human brucellosis in Maghreb: existence of a lineage related to socio-historical connections with Europe. *PLoS One* 2014;9:e115319.
21. **Foster JT, Walker FM, Rannals BD, Hussain MH, Drees KP et al.** African lineage *Brucella melitensis* isolates from Omani livestock. *Front Microbiol* 2018;8:2702.
22. **Pappas G, Papadimitriou P, Akritidis N, Christou L, Tsianos EV.** The new global map of human brucellosis. *Lancet Infect Dis* 2006;6:91–99.
23. **Rossetti CA, Arenas-Gamboia AM, Maurizio E.** Caprine brucellosis: a historically neglected disease with significant impact on public health. *PLoS Negl Trop Dis* 2017;11:e0005692.
24. **Musallam II, Abo-Shehada MN, Hegazy YM, Holt HR, Guitian FJ.** Systematic review of brucellosis in the middle East: disease frequency in ruminants and humans and risk factors for human infection. *Epidemiol Infect* 2016;144:671–685.
25. **European Food Safety Authority and European Centre for Disease Prevention and Control (EFSA and ECDC).** The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. *EFSA J* 2018;16:e05500.
26. **Corbel MJ.** *Brucellosis in Humans and Animals.* Geneva: World Health Organization; 2006.
27. **Dadar M, Shahali Y, Whatmore AM.** Human brucellosis caused by raw dairy products: a review on the occurrence, major risk factors and prevention. *Int J Food Microbiol* 2019;292:39–47.
28. **De Massis F, Di Girolamo A, Petrini A, Pizzigallo E, Giovannini A.** Correlation between animal and human brucellosis in Italy during the period 1997–2002. *Clin Microbiol Infect* 2005;11:632–636.
29. **Godfroid J.** Brucellosis in livestock and wildlife: zoonotic diseases without pandemic potential in need of innovative one health approaches. *Arch Public Health* 2017;75:34.
30. **European Commission.** 93/52/EEC: Commission decision of 21 December 1992 recording the compliance by certain member states or regions with the requirements relating to brucellosis (*B. melitensis*) and according them the status of a member state or region officially free of the disease. P. 14–15 (<https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A31993D0052>); 1992.
31. **De Massis F, Petrini A, Giovannini A.** Reliability evaluation of sampling plan fixed by Council Directive 91/68/EEC for the maintenance of officially brucellosis-free flock status. *J Vet Med B Infect Dis Vet Public Health* 2005;52:284–290.
32. **Graziani C, Mancini FR, Adone R, Maranelli C, Pasquali P et al.** La brucellosi in Italia dal 1998 al 2011. Istituto Superiore di Sanit, Roma 2013;Rapporti ISTISAN 13/45.
33. **Salute MD.** *Brucellosi - Piani Nazionali di Controllo 2018* (http://www.salute.gov.it/imgs/C_17_detttaglioPNI_1576_allegati_allegati_itemName_0_allegato.pdf) [accessed 2 April 2020]; 2018.
34. **De Massis F, Zilli K, Di Donato G, Nuvoloni R, Pelini S et al.** Distribution of *Brucella* field strains isolated from livestock, wildlife populations, and humans in Italy from 2007 to 2015. *PLoS One* 2019;14:e0213689.
35. **Salute MD.** *Bucellosi - Piani Nazionali di Controllo 2019* (<http://www.salute.gov.it/relazioneAnnuale2018/detttaglioRA2018.jsp?cap=capitolo1&sez=ra18-1-sanimale&id=2040>) [accessed 2 April 2020]; 2019.
36. **Facciola A, Palamara MAR, D'Andrea G, Marano F, Magliariti D et al.** Brucellosis is a public health problem in southern Italy: burden and epidemiological trend of human and animal disease. *J Infect Public Health* 2018;11:861–866.
37. **Janowicz A, De Massis F, Ancora M, Cammà C, Patavino C et al.** Core genome multilocus sequence typing and single nucleotide polymorphism analysis in the epidemiology of *Brucella melitensis* infections. *J Clin Microbiol* 2018;56:e00517–18.
38. **Sacchini L, Wahab T, Di Giannatale E, Zilli K, Abass A et al.** Whole genome sequencing for tracing geographical origin of imported cases of human brucellosis in Sweden. *Microorganisms* 2019;7:398.
39. **Andrews S.** FastQC. Babraham Bioinformatics, UK; 2010. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
40. **Bolger AM, Lohse M, Usadel B.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
41. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M et al.** SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
42. **Whatmore AM, Perrett LL, MacMillan AP.** Characterisation of the genetic diversity of *Brucella* by multilocus sequencing. *BMC Microbiol* 2007;7:34.
43. **Sahl JW, Beckstrom-Sternberg SM, Babic-Sternberg JS, Gillette JD, Hepp CM et al.** The In Silico Genotyper (ISG): an open-source pipeline to rapidly identify and annotate nucleotide variants for comparative genomics applications. *bioRxiv* 2015:015578.
44. **Li H, Durbin R.** Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754–1760.
45. **McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K et al.** The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–1303.
46. **Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A et al.** RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 2019;35:4453–4455.
47. **Letunic I, Bork P.** Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2019;47:W256–W259.
48. **Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J.** Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol Biol Evol* 2013;30:1224–1228.
49. **Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M et al.** BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 2019;15:e1006650.
50. **Menardo F, Loiseau C, Brites D, Coscolla M, Gygli SM et al.** Treemmer: a tool to reduce large phylogenetic datasets with minimal loss of diversity. *BMC Bioinformatics* 2018;19:164.
51. **Bouckaert RR, Drummond AJ.** bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evol Biol* 2017;17:42.
52. **Al Dahouk S, Neubauer H, Hensel A, Schöneberg I, Nöckler K et al.** Changing epidemiology of human brucellosis, Germany, 1962–2005. *Emerg Infect Dis* 2007;13:1895–1900.
53. **Georgi E, Walter MC, Pfalzgraf M-T, Northoff BH, Holdt LM et al.** Whole genome sequencing of *Brucella melitensis* isolated from 57 patients in Germany reveals high diversity in strains from Middle East. *PLoS One* 2017;12:e0175425.
54. **European Centre for Disease Prevention and Control.** *Surveillance atlas of infectious diseases* (<https://atlas.ecdc.europa.eu/public/index.aspx>) [accessed 7 September 2019]; 2019.
55. **Tan K-K, Tan Y-C, Chang L-Y, Lee KW, Nore SS et al.** Full genome SNP-based phylogenetic analysis reveals the origin and global spread of *Brucella melitensis*. *BMC Genomics* 2015;16:93.
56. **Cañón J, García D, García-Atance MA, Obexer-Ruff G, Lenstra JA et al.** Geographical partitioning of goat diversity in Europe and the Middle East. *Anim Genet* 2006;37:327–334.

57. Groeneveld LF, Lenstra JA, Eding H, Toro MA, Scherf B et al. Genetic diversity in farm animals - a review. *Anim Genet* 2010;41 (Suppl. 1):6–31.
58. Alberto FJ, Boyer F, Orozco-terWengel P, Streeter I, Servin B et al. Convergent genomic signatures of domestication in sheep and goats. *Nat Commun* 2018;9:813.
59. Lenstra JA, Econogene Consortium. Evolutionary and demographic history of sheep and goats suggested by nuclear, mtDNA and Y-chromosome markers. *Proceedings of the International Workshop – The Role of Biotechnology for the Characterization of Crop, Forestry, Animal and Fishery Genetic Resources, Turin, Italy, 5–7 March*; 2005.
60. Capasso L. Bacteria in two-millennia-old cheese, and related epizoonoses in Roman populations. *J Infect* 2002;45:122–127.
61. Kay GL, Sergeant MJ, Giuffra V, Bandiera P, Milanese M et al. Recovery of a medieval *Brucella melitensis* genome using shotgun metagenomics. *mBio* 2014;5:e01337-14.
62. Matteo S, Bellucci S. *Africa Italia: Due Continenti si Avvicinano*. Santarcangelo di Romagna: Fara Editore; 1999.

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.