



Extended Multilocus Sequence Analysis to Describe the Global Population Structure of the Genus *Brucella*: Phylogeography and Relationship to Biovars

Adrian M. Whatmore^{*}, Mark S. Koylass, Jakub Muchowski, James Edwards-Smallbone, Krishna K. Gopaul and Lorraine L. Perrett

FAO/WHO Collaborating Centre for Reference and Research in Brucellosis and OIE Brucellosis Reference Laboratory, Department of Bacteriology, Animal and Plant Health Agency, Addlestone, UK

OPEN ACCESS

Edited by:

Jorge Blanco, University of Santiago de Compostela, Spain

Reviewed by:

Vanesa García, University of Santiago de Compostela, Spain Ana Cristina Ferreira, Instituto Nacional de Investigação Agrária e Veterinária, Portugal

*Correspondence:

Adrian M. Whatmore adrian.whatmore@apha.gsi.gov.uk

Specialty section:

This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: 10 November 2016 Accepted: 06 December 2016 Published: 21 December 2016

Citation:

Whatmore AM, Koylass MS, Muchowski J, Edwards-Smallbone J, Gopaul KK and Perrett LL (2016) Extended Multilocus Sequence Analysis to Describe the Global Population Structure of the Genus Brucella: Phylogeography and Relationship to Biovars. Front. Microbiol. 7:2049. doi: 10.3389/fmicb.2016.02049 An extended multilocus sequence analysis (MLSA) scheme applicable to the Brucella, an expanding genus that includes zoonotic pathogens that severely impact animal and human health across large parts of the globe, was developed. The scheme, which extends a previously described nine locus scheme by examining sequences at 21 independent genetic loci in order to increase discriminatory power, was applied to a globally and temporally diverse collection of over 500 isolates representing all 12 known Brucella species providing an expanded and detailed understanding of the population genetic structure of the group. Over 100 sequence types (STs) were identified and analysis of data provided insights into both the global evolutionary history of the genus, suggesting that early emerging Brucella abortus lineages might be confined to Africa while some later lineages have spread worldwide, and further evidence of the existence of lineages with restricted host or geographical ranges. The relationship between biovar, long used as a crude epidemiological marker, and genotype was also examined and showed decreasing congruence in the order Brucella suis > B. abortus > Brucella melitensis. Both the previously described nine locus scheme and the extended 21 locus scheme have been made available at http://pubmlst.org/brucella/ to allow the community to interrogate existing data and compare with newly generated data.

Keywords: Brucella, brucellosis, multilocus sequence, molecular typing, zoonosis

INTRODUCTION

Brucellosis remains one of the world's most important zoonotic diseases and continues to have a significant impact on animal and human health in much of the world particularly in South America, Southern Europe, Africa, the Middle East, and much of Asia (Pappas et al., 2006). Recent years have seen the beginning of an expansion of the genus *Brucella* (Whatmore, 2009) from the six classically identified species [*Brucella abortus* (cattle), *Brucella melitensis* (sheep and goats), *Brucella suis* (pigs, hares, reindeer), *Brucella canis* (dogs), *B. ovis* (sheep), and *B. neotomae* (rodents)] with the description of six additional species (*Brucella microti* (voles), *Brucella pinnipedialis* (pinnipeds), *Brucella ceti* (cetaceans), *Brucella papionis* (baboons), *Brucella vulpis* (foxes), and *Brucella inopinata*

(isolated from a human case—natural host unknown). Although *Brucella* have long been known to represent a highly homogeneous genus, even for a time being classified as a monospecific genus (Verger et al., 1985), the known diversity of the group was extended greatly by the description of this latter species (De et al., 2008; Wattam et al., 2012) and *B. vulpis* (Hofer et al., 2012; Scholz et al., 2016). These groups are often described as "atypical" in the literature reflecting their genetic separation from classical species. A number of other isolates that await formal taxonomic description originating from sources as diverse as humans, frogs, fish, and additional rodents will likely extend diversity within both the classical group and newly emerging "atypical" *Brucella* in the near future (Tiller et al., 2010a,b; Godfroid et al., 2011; Eisenberg et al., 2012; Scholz and Vergnaud, 2013; Whatmore et al., 2015; Eisenberg et al., 2016).

Of the species described above B. melitensis, B. abortus, and B. suis are the most significant in terms of both animal and human disease impact. Where achieved, control has reflected a combination of measures including animal vaccination and/or control strategies ("test and slaughter") along with improved food hygiene standards. International trade standards are applied as part of efforts to help control spread of these pathogens (World Organisation for Animal Health, 2011). B. canis and B. ovis are less economically significant pathogens of animals and, of these species, only B. canis is known to cause rare infections of humans (Marzetti et al., 2013). The significance of Brucella infection in the marine ecosystem remains uncertain (Nymo et al., 2011) and there is only evidence of one genotype causing serious, but rarely reported, naturally acquired infections of humans (Whatmore et al., 2008). B. inopinata has been confined to a single human case with no reservoir of infection identified and, while B. microti appears highly virulent in its natural host and other rodents (Jiménez de Bagüés et al., 2010), any pathogenic potential for man remains to be elucidated.

The three major species, B. melitensis, B. abortus, and B. suis are divided into biovars by a biotyping scheme that for many years has been the gold standard for Brucella characterisation at both the species and subspecies level. The scheme, based on a combination of growth characteristics, biochemical reactions, serotyping and bacteriophage typing, distinguishes the six classical species and further subdivides the major species into seven (B. abortus), five (B. suis) or three (B. melitensis) biovars, respectively. The homogeneity of the classical species meant that slow progress was made in identifying molecular markers to consistently define species and particularly to type at the subspecies level. Thus, in spite of the obviously limited resolution offered, biovars became commonly used epidemiological markers with particular biovars associated with certain geographical areas or certain hosts. However biotyping in its traditional form is expensive, time-consuming, involves hazardous culture and its rather subjective nature means it needs to be carried out by highly experienced scientists and is likely prone to inconsistency between laboratories. Further this traditional method is likely to become increasingly less relevant as the genus expands and new species emerge that diverge from classical criteria. As more molecular diversity has become apparent with technological advances biotyping is increasingly being replaced by the use of frontline molecular tools notably various diagnostic PCRs based on genomic deletions or SNPs (Foster et al., 2008; Gopaul et al., 2008, 2010; López-Goñi et al., 2008). However the performance of such tools is critically dependent on their design being based on an accurate and comprehensive understanding of the population genetic structure of the groups they are designed to separate (Keim et al., 2004).

This paper updates current understanding of the population structure of *Brucella* based on an extended (from an existing nine locus scheme—Whatmore et al., 2007) and extensive multilocus sequence analysis (MLSA) examining 21 independent genetic loci applied to a geographically and temporally diverse collection of over 500 *Brucella* isolates. The data significantly add to understanding of the genetic diversity of the group providing insights into evolutionary history, phylogeography, the relationship between genotype and biovar and provides a comprehensive framework to further understanding of these issues, for the future placement of newly emerging or atypical isolates and for the rational design of robust and accurate rapid diagnostic assays.

METHODS

Bacteriology

Isolates were minimally cultured on serum dextrose agar and DNA preparations or crude lysates were prepared by standard procedures described previously (Whatmore et al., 2006) to serve as PCR template. Biotyping was undertaken by standard approaches as described elsewhere (Alton et al., 1988; Whatmore, 2009).

MLSA

A 21 locus MLSA scheme (BruMLSA21) was developed by identifying an additional 12 informative housekeeping gene loci and adding these to the BruMLSA9 scheme described earlier (Whatmore et al., 2007). Loci and corresponding primers used in the scheme are described in Table S1. These were designed such as that identical PCR procedures could be applied to all 21 loci and thus PCR parameters and downstream purification and sequencing procedures were all as described previously (Whatmore et al., 2007). BruMLSA21 profiles of 508 isolates representing all 12 Brucella species (172 B. abortus, 84 B. melitensis, 100 B. suis, 24 B. canis, 11 B. ovis, 73 B. ceti, 20 B. pinnipedialis, 7 B. microti, 3 B. neotomae, 2 B. papionis, 2 B. vulpis, 1 B. inopinata and 9 unclassified Brucella isolates) were obtained and used in the analysis described here. Sequences of two strains B. abortus 9-941 and B. abortus 2308 were extracted from previously published whole genome sequences (Chain et al., 2005; Halling et al., 2005).

Data Analysis

A representative strain of each genotype (sequence type or ST) was used for phylogenetic analysis. Sequences of the 21 loci were concatenated to produce a 10,257 bp sequence (including indels) for each genotype. Phylogenetic analysis was performed with the MEGA software, Version 5.2 (Tamura et al., 2011). Neighbour joining trees were constructed using the Jukes-Cantor model and

the percentage bootstrap confidence levels of internal branches were calculated from 1000 resamplings of the original data.

Minimum spanning trees were constructed in Bionumerics using the predefined template and the categorical coefficient. STs are represented by circles and the size of the circle is indicative of the number of isolates of that particular type. The coloring inside the circles indicates the *Brucella* species (**Figure 2**) or geographical origin (**Figures 4**, **6**, **8**). The different line types connecting genotypes reflect different numbers of shared loci as described in figure legends. The maximum neighbor difference used to create complexes indicated by the gray shading is given in individual figure legends.

PubMLST

In common with MLST/MLSA schemes for most other bacteria, an open and expandable database has been established at the PubMLST website using the Bacterial Isolate Genome Sequence Database (BIGSdb) platform (Jolley and Maiden, 2010). Databases containing allele descriptions and allelic profiles for both BruMLSA9 (Whatmore et al., 2007) and the extended BruMLSA21 described here, and a corresponding isolate database, are available at http://pubmlst.org/brucella/ where data can be interrogated and submissions of new data are encouraged. All BruMLSA21 data that forms the basis of the analyses described in this communication, as well a significant volume of previously undescribed BruMLSA9 data, have been deposited in these databases.

Other Molecular Testing

Confirmation of isolates as *B. abortus* vaccine strains S19 and RB51 was performed using previously described PCR assays (Sangari and Agüero, 1994; Vemulapalli et al., 1999) and confirmatory SNP based typing (Gopaul et al., 2010).

RESULTS AND DISCUSSION

Extension of Existing MLSA Scheme

A previously described MLSA scheme based on nine loci (BruMLSA9) provided limited resolution dividing 161 Brucella isolates into only 27 STs reflecting the relative genetic homogeneity of the classical Brucella species (Whatmore et al., 2007). In order to provide a supplementary scheme with increased resolution this scheme was extended by the addition of fragments representing 12 additional informative housekeeping genes to characterize 21 distinct loci in total. Here we describe its application to a temporally and geographically diverse collection of over 500 Brucella isolates representing all known species and biovars, and including all type strains, to describe the global population structure of the genus. The collection divided into 101 BruMLSA21 STs (Table S2) with the number of alleles varying from 10 in the case of *caiA* and *fbaA* to 27 in the case of *glk*, a gene included in BruMLSA9, and previously shown to be the most variable in this scheme. Strains of B. vulpis and the B. inopinatalike isolate BO2 failed to amplify a product at one locus (*mviM*) and were not assigned a BruMLSA21 ST although they have a full BruMLSA9 profile and ST. In addition the application of BruMLSA9 to a larger isolate collection than previously increased





the number of BruMLSA9 STs from 27 to 60 when applied to the strain collection described here (**Table S2**).

Overarching Population Structure

Individual allele sequences for each individual ST were concatenated to form a 10,257 bp length sequence for analysis with 493 (4.81%) sites being polymorphic excluding indels. With 219 polymorphic sites (2.14%) diversity is much reduced if only the core *Brucella* group are considered (i.e., without BruMLSA21 STs 57, 69, 70, and 101 representing *B. inopinata* and unclassified isolates from Australian rodents). Phylogenetic analysis based on all 101 STs is shown in **Figure S1** and illustrates the overall structure of the genus with *B. inopinata* and isolates from Australian rodents comprising early branching groups in the genus as currently described. Studies of a number of other groups including isolates from frogs (Eisenberg et al., 2012; Whatmore et al., 2015; Scholz et al., accepted) and foxes (*B.*

vulpis) (Scholz et al., 2016) for which BruMLSA9, or partial BruMLSA21, data are available confirm that these isolates also belong to the genetically atypical Brucella but that they are clearly much more closely affiliated to the genus Brucella than to the nearest phylogenetic neighbor Ochrobactrum. Figure 1 shows the core Brucella only as an expanded subtree to facilitate resolution. BruMLSA21 confirms the separation into clear clades comprising B. abortus/B. melitensis, B. suis/canis, B. ceti/B. pinnipedialis, B. ovis/B. papionis, and B. neotomae. These clades appear to have radiated virtually simultaneously from a common ancestor into host specific lineages. This was previously also seen with BruMLSA9 analysis and it was initially hoped that the additional discrimination offered by BruMLSA21 would help resolve this. However, it is now apparent that even whole genome sequencing phylogenies suggest the same explosive radiation has occurred although the reasons for this remain unclear (Audic et al., 2011; Wattam et al., 2014). Two additional short branches



represent *B. suis* biovar 5 (BruMLSA21 ST19) and *B. microti* the greater resolution offered by whole genome sequencing has suggested that the former is a very early branching group of the *B. suis/B. canis* lineage while the latter is just basal to the classical *Brucella* lineage (Audic et al., 2011; Wattam et al., 2014).

One additional lineage corresponds to BruMLSA21 STs 58 and 66—these represent historical isolates obtained by APHA in the 1960s reportedly associated with human infections in Thailand and originally characterized as atypical *B. suis*. These isolates might, with additional analysis, merit description as a novel species and illustrate the value of MLSA in placing novel isolates in the context of the complete extant understanding of the population structure of the genus. Such data adds to the evidence base facilitating description of novel species as has been the case for *B. papionis*, *B. inopinata*, *B. microti*, and *B. vulpis* (Scholz et al., 2008, 2010, 2016; Whatmore et al., 2014) and for other emerging atypical *Brucella* awaiting formal taxonomic description (Tiller et al., 2010a; Eisenberg et al., 2012; Whatmore et al., 2015).

Construction of a minimum spanning tree treating data categorically and defining clusters on the basis of neighbors differing in no more than five of the 21 loci revealed a similar relationship (**Figure 2**). Under these analysis conditions *B. melitensis* constituted a single complex, the more diverse *B. abortus* three complexes corresponding to Clades A, B, and C described below by intraspecies phylogenetic analysis, and *B. suis* biovars 1-4/B. *canis* two complexes, one corresponding to biovar 2 and the other biovars 1, 3, and 4 and *B. canis*. *B. papionis*, *B. ovis*, and *B. neotomae* all constitute separate groups

while the remaining major central complex contains *B. ceti. B. pinnipedialis, B. microti,* and *B. suis* biovar 5 isolates.

Intraspecies Relationships within *B. Abortus:* Phylogeography and Biovar

Only six *B. abortus* STs were described initially by BruMLSA9 application of BruMLSA21 to a wider strain collection significantly increased known diversity of *B. abortus* with 30 STs identified. Isolates comprise three major clades A, B, and C (**Figures 3**, **4**)—the two early branching clades A and B comprise isolates entirely originating from widely across Africa (Senegal, Nigeria, Zimbabwe, Sudan, Mozambique, Kenya, Chad, and Uganda) with clade A representing a previously undescribed clade that substantially increases known diversity of *B. abortus*. These clades consist predominantly isolates of biovars 1, 3, and 6 including the biovar 3 reference strain Tulya.

In contrast to the geographical restriction of clades A and B, isolates of clade C have a global distribution with in particular ST1 and ST2 appearing to be widely distributed across many

continents. As shown in Figure 3 this clade splits into two subclades-in clade C1 the most common genotype to biovar association is with biovar 3. However one ST, ST30 is associated exclusively with isolates of biovars 4, 5, and 9 and includes virtually all isolates of the latter two biovars and their reference strains. The vast majority of these isolates originate from the UK (prior to brucellosis eradication) with most of the few remaining isolates from other parts of Western Europe. Consistent with this biovar 5 was colloquially known as "British melitensis" (Corbel and Banai, 2005) and is differentiated from biovar 9, colloquially known as "H₂S producing melitensis" (Corbel and Banai, 2005), only by H₂S production. The remaining isolates of ST30 represent biovar 4 which is also seen in genetically distinct lineages and differs from biovar 5 and 9 only it its inability to grow in the presence of thionin. While the overall picture within Clade C is of global dispersal there is evidence that individual genotypes may reflect types with restricted geographical range that may be endemic in certain regions-notably STs 63 and 31 were exclusively associated with East Asia (India, Sri Lanka, and





Thailand). Much more detailed local analysis, enabled by the work described here, will help better understand the genotypes prevalent at local levels.

Clade C2 appears more strongly associated with biovar 1 isolates. The particularly widely dispersed genotype, ST1, consists of isolates of biovar 1, 2, and 4 including the reference strains for all these biovars. Biovars 1 and 2 differ only in fuchsin susceptibility while biovar 4 isolates show a different pattern of agglutination with monospecific sera. *B. abortus* biovars 1, 2, and 4 have been shown historically to be closely genetically related (e.g., Gargani and López-Merino, 2006) and, based on these findings it is still not possible to identify distinct genetic lineages corresponding to these biovars.

In summary these data, which reflect the first extensive genetic characterisation of *Brucella* from diverse African origins, suggest that *B. abortus* originated in Africa and, the fact that isolates of the two earliest emerging branches (clades A and B) are still confined there to date, may suggest no or limited global spread of these lineages. In contrast the later emerging clade C appears to have been spread globally to all six populated continents. Indeed it is interesting to note that sporadic members of this clade have been isolated from Africa which may suggest these newer lineages were introduced to Africa by cattle importation from Europe such that both the original endemic lineages (Clades A and B) and, probably to a much lesser extent, introduced lineages (Clades C1 and C2) may now circulate in Africa.

While some relationship between genotype and biovar is apparent in *B. abortus* the picture is far from one of complete

congruence. For example isolates of biovar 1 are widely distributed in clades B and C2 and, to a lesser extent in C1. Most isolates in the major early diverging African clade correspond to biovars 1, 3, and 6 (with the latter two biovars separable only by the requirement of biovar 3 for additional CO2 and use of thionin at 1/25,000 dilution) although, of the reference strains, only that of biovar 3 is found in this clade. Notably clade C1 is also strongly associated with biovar 3-these data appear consistent with reports in the literature of two biovar 3 clades, 3a from Africa (our Clade B) and 3b (our Clade C1) from Europe (Ocampo-Sosa et al., 2005; Ica et al., 2008; Bertu et al., 2015). Biovar 3b isolates have been reported from Spain and Turkey consistent with the sources of isolates falling within clade C1 (Table S2). In contrast, biovars 5 and 9 were almost entirely confined to a single ST within clade C1 suggesting these biovars are congruent with a distinct genetic lineage although clearly genetic separation between these two biovars was not possible using this MLSA scheme. The biovar 6 reference strain is also located in clade C1 albeit with a unique genotype, ST4, consistent with the description of a PCR assay specific for *B. abortus* biovars 3b, 5, 6, and 9 although this notably was only applied to reference strains of the latter three biovars (Ocampo-Sosa et al., 2005). Readers should note that the original biovar 7 reference strain (63/75) has been shown to represent a mixed culture and biovar 7 was thus deleted from Brucella taxonomy (Corbel, 1988; Garin-Bastuji et al., 2014). However genetically diverse isolates that possess the phenotypic profile reported for biovar 7 have been included here and reported in other studies (Garin-Bastuji et al., 2014; Kim et al., 2016). The genetic diversity of strains with the biovar 7 phenotype questions any move to reintroduce biovar 7 as previously defined. Of the remaining biovars, biovar 2 is almost exclusively associated with ST1 within clade C2 that also includes the majority of biovar 4 strains and a large number of biovar 1 strains. The reference strains for all three biovars 1, 2, and 4 are members of ST1 highlighting the fact that it not possible to identify defined and unique genetic lineages that correspond to these biovars.

Finally both *B. abortus* vaccine strains S19 and RB51 were found to be members of ST5 consistent with the identification of North American field isolates in this ST (the parental strains of the vaccines were North American isolates). Given the sharing of this ST between field and vaccine strains ST5 strains were examined by molecular assays to determine status as vaccine or field strains (Sangari and Agüero, 1994; Vemulapalli et al., 1999; Gopaul et al., 2010). These analyses identified that a substantial proportion of isolates in this ST appear to represent S19 vaccine re-isolated from the field (**Table S2**).

Intraspecies Relationships within *B. melitensis*: Phylogeography and Biovar

Application of BruMLSA21 to a wider *B. melitensis* strain collection identified 21 STs in contrast to the 6 STs previously described by BruMLSA9. Isolates cluster into three distinct lineages (**Figures 5, 6**) that correspond to the "Americas," "West Mediterranean," and "East Mediterranean" lineages described previously on the basis of MLVA data (Al Dahouk et al., 2007). Only a small number of isolates from the Americas were examined here and while isolates from USA and Argentina are included in the Americas clade more than 60% of the isolates of this clade are of African origin. African isolates included in this clade originate widely across the continent (Ethiopia, Somalia, Nigeria, Tanzania, Sudan, Zimbabwe) and, as few African isolates are seen in the "West Mediterranean" and "East Mediterranean" lineages, it appears that this lineage may be endemic in Africa.

In agreement with the previous designation isolates falling in the "West Mediterranean" lineage were exclusively reported from Italy, France, the former Yugoslavia and Germany. In contrast



the "East Mediterranean" lineage, while including isolates from Greece, Turkey, Cyprus, and the Balkan States, also extends much more widely into the Middle East and Asia (Thailand, India, Pakistan, Mongolia, and Afghanistan). Although numbers are small some BruMLSA21 STs appear geographically restricted—for example ST82 associated with Thailand (and in agreement with the identification by a recent whole genome sequence analysis of a separate lineage within the "East Mediterranean" cluster including the ST82 defining strain and strains from Malaysia and the Philippines—Tan et al., 2015) and ST38 to the Middle East. However ST8, by far the most frequently seen *B. melitensis* ST, appears widely distributed across Europe and Asia from Portugal to Afghanistan and Mongolia.

While each of the three lineages contains one of the *B. melitensis* biovar reference strains ("West Mediterranean" Ether biovar 3; "East Mediterranean" 63/9 biovar 2; "Americas" 16M biovar 1) there is no clear relationship between genotype and biovar. While the majority of isolates in the "Americas" lineage belong to biovar 1 (81%) and in the "West Mediterranean" lineage belong to biovar 3 (64%), isolates of both other biovars were apparent in both of these lineages. In addition virtually equivalent numbers of all three biovars are seen in both the "East Mediterranean" lineage overall and the most common

and widely dispersed BruMLSA21 ST8 specifically. Early MLVA studies highlighted a lack of congruence between genotype and biovar (Whatmore et al., 2006) and a growing body later MLVA studies (e.g., Al Dahouk et al., 2007; Jiang et al., 2011, 2013), taken together with data presented here using markers more appropriate for determining phylogeny, suggest that the biovar concept, which is the case of *B. melitensis* is reliant solely on serological reaction, is of extremely limited epidemiological value for *B. melitensis*.

Intraspecies Relationships of *B. suis/B. canis* Lineage: Phylogeography and Biovar

As with other species application of BruMLSA21 to a wider *B. suis* strain collection significantly increased the number of STs identified in BruMLSA9 from 6 to 20. In contrast to *B. abortus* and *B. melitensis* there is a strong congruency between genotype and biovar. Excluding *B. suis* biovar 5 there are two genetically divergent major clades (**Figure 7**), one comprising to biovar 2 isolates and the second comprising an early branching group corresponding to biovar 1 isolates, biovar 3 and 4 isolates and *B. canis* isolates forming a terminal group on this branch. Thus only



Frontiers in Microbiology | www.frontiersin.org

biovar 3 and 4 isolates do not represent clearly separated clusters although, in contrast to BruMLSA9 where they share a genotype, the small number of biovar 3 isolates included do correspond to a single unique BruMLSA21 ST (ST17).

Only biovar 2 has sufficient isolates to ascertain possible phylogeographic associations (**Figure 8**). Here the vast majority of STs are associated with Central and Northern European isolates where it is likely that movement of strains in wildlife reservoirs (hares and wild boar) has dispersed genotypes widely. However two genotypes ST39 and ST44 are confined to the Iberian peninsula suggesting they correspond to the Iberian clone of *B. suis* biovar 2 described previously on the basis of PCR-RFLP, MLVA and subsequent whole genome mapping (Ferreira et al., 2016).

Intraspecies Relationships of *B. pinnipedialis* and *B. ceti*

Grouping of marine mammal *Brucella* remains largely consistent with that described by BruMLSA9 with an increase in STs from 5 to 16 (**Figure 9** and **Table S2**). As suggested previously phylogeny appears inconsistent with taxonomy in this group and *B. ceti* appears paraphyletic with two distinct clusters (ST26 complex and ST23 complex) that appear to have preferred hosts of dolphins or porpoises, respectively (Groussaud et al., 2007; Dawson et al., 2008). The increased resolution of BruMLSA21 allows further subdivision of isolates of BruMLSA9 ST24 and ST25 associated with various species of seals for example identifying two BruMLSA21 STs (ST53 and ST54) that appear to be specific for hooded seals in this dataset, a finding consistent with the previous description of a hooded seal specific clade on the basis of MLVA (Groussaud et al., 2007; Maquart et al., 2009). The remaining ST, ST27 has previously been associated with human zoonotic infection (Whatmore et al., 2008) and remains distinct from the main marine mammal *Brucella* clusters. These data again highlight that phylogenetic relationships suggest that a taxonomic rearrangement of *Brucella* from marine mammals might be appropriate (Whatmore, 2009).

Remaining Species

No diversity was identified in *B. microti, B. neotomae, B. papionis,* and *B. ovis.* In large part this may reflect the very limited sampling of these rarely or recently isolated species although *B. ovis* has been more extensively sampled with isolates from across the globe included here. These data, and complementary MLVA





data, which also reveals relatively limited diversity within *B. ovis* (Whatmore et al., 2006), suggest that *B. ovis*, while relatively well-separated from other classical species, may represent an extremely genetically restricted group.

Only a single isolate of *B. inopinata* has been formally described although the increasing isolation of diverse genetically atypical *Brucella* from various hosts might point toward consideration of whether the description of *B. inopinata* should be expanded to include all these early-branching *Brucella* isolates.

CONCLUSIONS

Although we have previously described the application of BruMLSA21 to address specific issues, for example in the context of addressing the status of *B. abortus* biovar 7 (Garin-Bastuji et al., 2014), full technical details of the scheme, allele calls, ST allelic profiles, and associated metadata have not previously been available to facilitate comparative analysis by others.

To bring *Brucella* into alignment with most other bacterial pathogens the existing BruMLSA9 data (Whatmore et al., 2007) and BruMLSA21 data described here, as well as a large volume of additional unpublished BruMLSA9 data has now been made available via the principle MLST/MLSA database as a resource for the scientific community. The full data set, sampling c. over 740 isolates for BruMLSA9 and more than 500 isolates for BruMLSA21, provide the most comprehensive assessment of the population structure

of the entire genus to date. They provide a framework for understanding the relationship between extant strains, for the placement of newly emerging or discovered strains in the context of extant knowledge and a platform for designing and or assessing the performance of molecular diagnostic assays designed to characterize strains at various taxonomic levels.

The study described here adds to understanding of the phylogeographical relationships within the major Brucella species for the first time clearly describing the apparent origin of B. abortus in Africa and the global spread of later emerging lineages. Completely novel lineages, such as the early emerging African B. abortus lineage, have been revealed which may be significant in understanding the emergence of *B. abortus* as a host specific pathogen. Equally there appear to be clones endemic only locally, such as the B. abortus ST63/ST31 Asian clone or the Middle Eastern B. melitensis ST38, although much of the world, particularly where brucellosis is most problematic, remains grossly under sampled with insufficient data to draw robust conclusions. It is hoped the publication of these data encourage the application of MLSA (either perse or using data extracted from whole genome sequences) more widely to expand the depth of the database described here and further understanding of the global Brucella situation.

The data also add significantly to understanding the relationship between biovar and genotype—it is clear that there is there is diminishing congruence in this relationship

from B. suis > B. abortus > B. melitensis. This is not entirely unexpected given a system based on phenotypic or serological markers that may be impacted by factors such as variable expression and the acknowledged subjectivity of the approach combined with the very subtle differences in phenotype of some of the biovars described. Given advances in molecular approaches it would seem timely to consider whether such approaches offer a more meaningful replacement for biotyping. Such assays could be based on genomic rearrangements such as the multiplex Bruceladder (López-Goñi et al., 2008) or SNPs as per existing speciation assays (Scott et al., 2007; Foster et al., 2008; Gopaul et al., 2008) which have the advantage that they can be applied at a number of taxonomic levels (Keim et al., 2004). The crucial point is that they must be designed on the back of a robust population genetic understanding provided by data such as that presented here and iteratively re-examined as knowledge of extant diversity increases. For example it would be very easy to design an assay for B. abortus that did not detect the early emerging African lineage if a marker was selected that represented an evolutionary change that occurred beyond this branch.

We welcome submission of new data to the *Brucella* PubMLST website at http://pubmlst.org/brucella/.

AUTHORS CONTRIBUTIONS

AW conceived of and designed the study, analyzed the data, and drafted the manuscript. MK, JM, KG, and JE undertook

REFERENCES

- Al Dahouk, S., Flèche, P. L., Nöckler, K., Jacques, I., Grayon, M., Scholz, H. C., et al. (2007). Evaluation of *Brucella* MLVA typing for human brucellosis. *J. Microbiol. Methods* 69, 137–145. doi: 10.1016/j.mimet.2006.12.015
- Alton, G. G., Jones, L. M., Angus, R. D., and Verger, J. M. (1988). Techniques for the Brucellosis Laboratory. Paris: Institut National de la Recherche Agronomique.
- Audic, S., Lescot, M., Claverie, J. M., Cloeckaert, A., and Zygmunt, M. S. (2011). The genome sequence of *Brucella pinnipedialis* B2/94 sheds light on the evolutionary history of the genus *Brucella*. *BMC Evol. Biol.* 11:200. doi: 10.1186/1471-2148-11-200
- Bertu, W. J., Ducrotoy, M. J., Muñoz, P. M., Mick, V., Zúñiga-Ripa, A., Bryssinckx, W., et al. (2015). Phenotypic and genotypic characterization of *Brucella* strains isolated from autochthonous livestock reveals the dominance of *B. abortus* biovar 3a in Nigeria. *Vet. Microbiol.* 180, 103–108. doi: 10.1016/j.vetmic.2015.08.014
- Chain, P. S., Comerci, D. J., Tolmasky, M. E., Larimer, F. W., Malfatti, S. A., Vergez, L. M., et al. (2005). Whole-genome analyses of speciation events in pathogenic *Brucellae*. *Infect. Immun.* 73, 8353–8361. doi: 10.1128/IAI.73.12.8353-8361.2005
- Corbel, M. J., and Banai, M. (2005). "Genus I. Brucella Meyer and Shaw 1920, 173AL," in Bergey's Manual of Systematic Bacteriology, Vol. 2, eds D. J. Brenner, N. R. Krieg, and J. T. Staley (New York, NY: Springer), 370–386.
- Corbel, M. J. (1988). International committee on systematic bacteriology, subcommittee on the taxonomy of *Brucella. Int. J. Syst. Bacteriol.* 38, 450–452. doi: 10.1099/00207713-38-4-450
- Dawson, C. E., Stubberfield, E. J., Perrett, L. L., King, A. C., Whatmore, A. M., Bashiruddin, J. B., et al. (2008). Phenotypic and molecular characterisation of *Brucella* isolates from marine mammals. *BMC Microbiol.* 8:224. doi: 10.1186/1471-2180-8-224

portions of the experimental work and analysis. LP oversaw strain provision and biotyping.

ACKNOWLEDGMENTS

Brucellosis research at the APHA is supported by the UK Department of the Environment, Food and Rural Affairs (Defra). Development and implementation of MLSA typing was specifically funded as part of projects SE0311 and SE0313 with ongoing support from Defra for database maintenance and expansion. We acknowledge the many colleagues worldwide who have submitted strains to the APHA reference laboratory over the years facilitating the development of an extensive strain collection to support these studies.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb. 2016.02049/full#supplementary-material

Figure S1 | Phylogenetic reconstruction of relationships between all 101 BruMLSA21 STs including atypical strains. The tree was constructed with the concatenated sequence data of the 21 loci (>10.2 kb) using the neighbor-joining algorithm with the Jukes-Cantor model in MEGA5.2. Bar = nucleotide substitutions per site.

Table S1 | Technical details of MLSA scheme including details of loci characterized and PCR primer sequences.

Table S2 | Details of BruMLSA21 data described in this study. All data have been deposited at http://pubmlst.org/brucella/.

- De, B. K., Stauffer, L., Koylass, M. S., Sharp, S. E., Gee, J. E., Helsel, L. O., et al. (2008). Novel *Brucella* strain (BO1) associated with a prosthetic breast implant infection. *J. Clin. Microbiol.* 46, 43–49. doi: 10.1128/JCM.01 494-07
- Eisenberg, T., Hamann, H. P., Kaim, U., Schlez, K., Seeger, H., Schauerte, N., et al. (2012). Isolation of potentially novel *Brucella* spp. from frogs. *Appl. Environ. Microbiol.* 78, 3753–3755. doi: 10.1128/AEM. 07509-11
- Eisenberg, T., Riße, K., Schauerte, N., Geiger, C., Blom, J., and Scholz, H. C. (2016). Isolation of a novel "atypical" *Brucella* strain from a bluespotted ribbontail ray (*Taeniura lymma*). *Antonie Van Leeuwenhoek*. doi: 10.1007/s10482-016-0792-4. [Epub ahead of print].
- Ferreira, A. C., Dias, R., de Sá, M. I., and Tenreiro, R. (2016). Whole-genome mapping reveals a large chromosomal inversion on Iberian *Brucella suis* biovar 2 strains. *Vet. Microbiol.* 192, 220–225. doi: 10.1016/j.vetmic.2016.07.024
- Foster, J. T., Okinaka, R. T., Svensson, R., Shaw, K., De, B. K., Robison, R. A., et al. (2008). Real-time PCR assays of single-nucleotide polymorphisms defining the major *Brucella* clades. *J. Clin. Microbiol.* 46, 296–301. doi: 10.1128/JCM.01496-07
- Gargani, G., and López-Merino, A. (2006). International Committee on Systematic Bacteriology; Subcommittee on the taxonomy of *Brucella. Int. J. Syst. Evol. Microbiol.* 56, 1167–1168. doi: 10.1099/ijs.0.64042-0
- Garin-Bastuji, B., Mick, V., Le Carrou, G., Allix, S., Perrett, L. L., Dawson, C. E., et al. (2014). Examination of taxonomic uncertainties surrounding *Brucella abortus* bv. 7 by phenotypic and molecular approaches. *Appl. Environ. Microbiol.* 80, 1570–1579. doi: 10.1128/AEM.03755-13
- Godfroid, J., Scholz, H. C., Barbier, T., Nicolas, C., Wattiau, P., Fretin, D., et al. (2011). Brucellosis at the animal/ecosystem/human interface at the beginning of the 21st century. *Prev. Vet. Med.* 102, 118–131. doi: 10.1016/j.prevetmed.2011.04.007

- Gopaul, K. K., Koylass, M. S., Smith, C. J., and Whatmore, A. M. (2008). Rapid identification of *Brucella* isolates to the species level by real time PCR based single nucleotide polymorphism (SNP) analysis. *BMC Microbiol.* 8:86. doi: 10.1186/1471-2180-8-86
- Gopaul, K. K., Sells, J., Bricker, B. J., Crasta, O. R., and Whatmore, A. M. (2010). Rapid and reliable single nucleotide polymorphism-based differentiation of *Brucella* live vaccine strains from field strains. *J. Clin. Microbiol.* 48, 1461–1464. doi: 10.1128/JCM.02193-09
- Groussaud, P., Shankster, S. J., Koylass, M. S., and Whatmore, A. M. (2007). Molecular typing divides marine mammal strains of *Brucella* into at least three groups with distinct host preferences. *J. Med. Microbiol.* 56, 1512–1518. doi: 10.1099/jmm.0.47330-0
- Halling, S. M., Peterson-Burch, B. D., Bricker, B. J., Zuerner, R. L., Qing, Z., Li, L. L., et al. (2005). Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. J. Bacteriol. 187, 2715–2726. doi: 10.1128/JB.187.8.2715-27 26.2005
- Hofer, E., Revilla-Fernández, S., Al Dahouk, S., Riehm, J. M., Nöckler, K., Zygmunt, M. S., et al. (2012). A potential novel *Brucella* species isolated from mandibular lymph nodes of red foxes in Austria. *Vet. Microbiol.* 155, 93–99. doi: 10.1016/j.vetmic.2011.08.009
- Ica, T., Aydin, F., Erdenlig, S., Guler, L., and Büyükcangaz, E. (2008). Characterisation of *Brucella abortus* biovar 3 isolates from Turkey as biovar 3b. Vet. Rec. 163, 659–661. doi: 10.1136/vr.163.22.659
- Jiang, H., Fan, M., Chen, J., Mi, J., Yu, R., Zhao, H., et al. (2011). MLVA genotyping of Chinese human *Brucella melitensis* biovar 1, 2 and 3 isolates. *BMC Microbiol*. 11:256. doi: 10.1186/1471-2180-11-256
- Jiang, H., Wang, H., Xu, L., Hu, G., Ma, J., Xiao, P., et al. (2013). MLVA genotyping of *Brucella melitensis* and *Brucella abortus* isolates from different animal species and humans and identification of *Brucella suis* vaccine strain S2 from cattle in China. *PLoS ONE* 8:e76332. doi: 10.1371/journal.pone.0076332
- Jiménez de Bagüés, M. P., Ouahrani-Bettache, S., Quintana, J. F., Mitjana, O., Hanna, N., Bessoles, S., et al. (2010). The new species *Brucella microti* replicates in macrophages and causes death in murine models of infection. *J. Infect. Dis.* 202, 3–10. doi: 10.1086/653084
- Jolley, K. A., and Maiden, M. C. (2010). BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11:595. doi: 10.1186/1471-2105-11-595
- Keim, P., Van Ert, M. N., Pearson, T., Vogler, A. J., Huynh, L. Y., and Wagner, D. M. (2004). Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. *Infect. Genet. Evol.* 4, 205–213. doi: 10.1016/j.meegid.2004.02.005
- Kim, J.-Y., Kang, S.-I., Erdenebaataar, J., Vanaabaatar, B., Ulziisaikhan, G., Khursbaatar, O., et al. (2016). Distinctive phenotypic and molecular characteristics of *Brucella abortus* strains isolated from Mongolia. *Turk. J. Vet. Anim. Sci.* 40, 562–568. doi: 10.3906/vet-1511-75
- López-Goñi, I., García-Yoldi, D., Marín, C. M., de Miguel, M. J., Muñoz, P. M., Blasco, J. M., et al. (2008). Evaluation of a multiplex PCR assay (Bruce-ladder) for molecular typing of all *Brucella* species, including the vaccine strains. *J. Clin. Microbiol.* 46, 3484–3487. doi: 10.1128/JCM.00837-08
- Maquart, M., Le Flèche, P., Foster, G., Tryland, M., Ramisse, F., Djønne, B., et al. (2009). MLVA-16 typing of 295 marine mammal *Brucella* isolates from different animal and geographic origins identifies 7 major groups within *Brucella ceti* and *Brucella pinnipedialis*. *BMC Microbiol*. 9:145. doi: 10.1186/1471-2180-9-145
- Marzetti, S., Carranza, C., Roncallo, M., Escobar, G. I., and Lucero, N. E. (2013). Recent trends in human *Brucella canis* infection. *Comp. Immunol. Microbiol. Infect. Dis.* 36, 55–61. doi: 10.1016/j.cimid.2012.09.002
- Nymo, I. H., Tryland, M., and Godfroid, J. (2011). A review of *Brucella* infection in marine mammals with special emphasis on *Brucella pinnipedialis* in the hooded seal (*Cystophora cristata*). Vet. Res. 42:93. doi: 10.1186/1297-9716-42-93
- Ocampo-Sosa, A. A., Agüero-Balbín, J., and García-Lobo, J. M. (2005). Development of a new PCR assay to identify *Brucella abortus* biovars 5, 6 and 9 and the new subgroup 3b of biovar 3. *Vet. Microbiol.* 110, 41–51. doi: 10.1016/j.vetmic.2005.06.007
- Pappas, G., Papadimitriou, P., Akritidis, N., Christou, L., and Tsianos, E. V. (2006). The new global map of human brucellosis. *Lancet Infect. Dis.* 6, 91–99. doi: 10.1016/S1473-3099(06)70382-6

- Sangari, F. J., and Agüero, J. (1994). Identification of *Brucella abortus* B19 vaccine strain by the detection of DNA polymorphism at the ery locus. *Vaccine* 12, 435–438. doi: 10.1016/0264-410X(94)90121-X
- Scholz, H. C., and Vergnaud, G. (2013). Molecular characterisation of *Brucella* species. *Rev. Sci. Tech.* 32, 149–162. doi: 10.20506/rst.32.1.2189
- Scholz, H. C., Hubalek, Z., Sedlácek, I., Vergnaud, G., Tomaso, H., Al Dahouk, S., et al. (2008). *Brucella microti* sp. nov., isolated from the common vole *Microtus arvalis. Int. J. Syst. Evol. Microbiol.* 58, 375–382. doi: 10.1099/ijs.0. 65356-0
- Scholz, H. C., Nöckler, K., Göllner, C., Bahn, P., Vergnaud, G., Tomaso, H., et al. (2010). *Brucella inopinata* sp. nov., isolated from a breast implant infection. *Int. J. Syst. Evol. Microbiol.* 60, 801–808. doi: 10.1099/ijs.0. 011148-0
- Scholz, H. C., Revilla-Fernández, S., Al Dahouk, S., Hammerl, J. A., Zygmunt, M. S., Cloeckaert, A., et al. (2016). Brucella vulpis sp. nov., isolated from mandibular lymph nodes of red foxes (Vulpes vulpes). Int. J. Syst. Evol. Microbiol. 66, 2090–2098. doi: 10.1099/ijsem.0. 000998
- Scholz, H. C., Muhldorfer, K., Shilton, C., Bendict, S., Whatmore, A. M., and Eisenberg, T. (accepted). Change of a medically important genus: Worldwide emergence of genetically diverse novel *Brucella* species in exotic frogs. *PLoS ONE*.
- Scott, J. C., Koylass, M. S., Stubberfield, M. R., and Whatmore, A. M. (2007). Multiplex assay based on single-nucleotide polymorphisms for rapid identification of *Brucella* isolates at the species level. *Appl. Environ. Microbiol.* 73, 7331–7337. doi: 10.1128/AEM.00976-07
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121
- Tan, K. K., Tan, Y. C., Chang, L. Y., Lee, K. W., Nore, S. S., Yee, W. Y., et al. (2015). Full genome SNP-based phylogenetic analysis reveals the origin and global spread of *Brucella melitensis*. *BMC Genomics*. 16:93. doi: 10.1186/s12864-015-1294-x
- Tiller, R. V., Gee, J. E., Frace, M. A., Taylor, T. K., Setubal, J. C., Hoffmaster, A. R., et al. (2010a). Characterization of novel *Brucella* strains originating from wild native rodent species in North Queensland, Australia. *Appl. Environ. Microbiol.* 76, 5837–5845. doi: 10.1128/AEM. 00620-10
- Tiller, R. V., Gee, J. E., Lonsway, D. R., Gribble, S., Bell, S. C., Jennison, A. V., et al. (2010b). Identification of an unusual *Brucella* strain (BO2) from a lung biopsy in a 52 year-old patient with chronic destructive pneumonia. *BMC Microbiol*. 10:23. doi: 10.1186/1471-2180-10-23
- Vemulapalli, R., McQuiston, J. R., Schurig, G. G., Sriranganathan, N., Halling, S. M., and Boyle, S. M. (1999). Identification of an IS711 element interrupting the *wboA* gene of *Brucella abortus* vaccine strain RB51 and a PCR assay to distinguish strain RB51 from other *Brucella* species and strains. *Clin. Diagn. Lab. Immunol.* 6, 760–764.
- Verger, J. M., Grimont, F., Grimont, P. A. D., and Grayon, M. (1985). Brucella, a monospecific genus as shown by deoxyribonucleic acid hybridization. Int. J. Syst. Bacteriol. 35, 292–295. doi: 10.1099/00207713-35-3-292
- Wattam, A. R., Inzana, T. J., Williams, K. P., Mane, S. P., Shukla, M., Almeida, N. F., et al. (2012). Comparative genomics of early-diverging *Brucella* strains reveals a novel lipopolysaccharide biosynthesis pathway. *Mbio* 3:e00246–212. doi: 10.1128/mBio.00388-12
- Wattam, A. R., Foster, J. T., Mane, S. P., Beckstrom-Sternberg, S. M., Beckstrom-Sternberg, J. M., Dickerman, A. W., et al. (2014). Comparative phylogenomics and evolution of the Brucellae reveal a path to virulence. J. Bacteriol. 196, 920–930. doi: 10.1128/JB.01091-13
- Whatmore, A. M., Shankster, S. J., Perrett, L. L., Murphy, T. J., Brew, S. D., Thirlwall, R. E., et al. (2006). Identification and characterization of variable number of tandem repeat markers for typing of *Brucella* spp. *J. Clin. Microbiol.* 44, 1982–1993. doi: 10.1128/JCM.02039-05
- Whatmore, A. M., Perrett, L. L., and MacMillan, A. P. (2007). Characterisation of the genetic diversity of *Brucella* by multilocus sequencing. *BMC Microbiol*. 7:34. doi: 10.1186/1471-2180-7-34
- Whatmore, A. M., Dawson, C. E., Groussaud, P., Koylass, M. S., King, A. C., Shankster, S. J., et al. (2008). Marine mammal *Brucella* genotype

associated with zoonotic infection. *Emerg. Infect. Dis.* 14, 517–518. doi: 10.3201/eid1403.070829

- Whatmore, A. M., Davison, N., Cloeckaert, A., Al Dahouk, S., Zygmunt, M. S., Brew, S. D., et al. (2014). *Brucella papionis* sp. nov., isolated from baboons (*Papio* spp.). *Int. J. Syst. Evol. Microbiol.* 64, 4120–4128. doi: 10.1099/ijs.0.065482-0
- Whatmore, A. M., Dale, E.-J., Stubberfield, E., Muchowski, J., Koylass, M., Dawson, C., et al. (2015). Isolation of *Brucella* from a White's tree frog (*Litoria caerulea*). *JMM Case Rep.* 2015:2. doi: 10.1099/jmmcr.0.000017
- Whatmore, A. M. (2009). Current understanding of the genetic diversity of Brucella, an expanding genus of zoonotic pathogens. Infect. Genet. Evol. 9, 1168–1184. doi: 10.1016/j.meegid.2009.07.001
- World Organisation for Animal Health (2011). Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. Available online from: http://www.oie.int/ international-standard-setting/terrestrial-manual/access-online/

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer VG and handling Editor declared their shared affiliation and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Whatmore, Koylass, Muchowski, Edwards-Smallbone, Gopaul and Perrett. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.