Genomic Diversity and Zoonotic Potential of *Brucella neotomae*

Gilles Vergnaud, Michel S. Zygmunt, Roland T. Ashford, Adrian M. Whatmore, Axel Cloeckaert

After reports in 2017 of *Brucella neotomae* infections among humans in Costa Rica, we sequenced 12 strains isolated from rodents during 1955–1964 from Utah, USA. We observed an exact strain match between the human isolates and 1 Utah isolate. Independent confirmation is required to clarify *B. neotomae* zoonotic potential.

The genus *Brucella* comprises a monophyletic group including 6 classical species showing clonal evolution: *B. abortus*, *B. suis*, *B. melitensis*, *B. canis*, *B. ovis*, *and B. neotomae* (*1*,*2*). The zoonotic potential of *B. melitensis*, *B. abortus*, *B. suis*, and *B. canis* (in decreasing order of disease burden in human populations) has been clinically established on the basis of numerous human cases reported over the past century.

B. neotomae was originally isolated from a single rodent species (desert woodrat, *Neotoma lepida*), in an area with low population density of other wild animals and remote from domestic livestock (*3*). Recently, 2 publications described the isolation in Costa Rica of *B. neotomae* strains from 2 human patients with brucellosis (*4*,*5*). According to those reports, the 2 human isolates, bneohCR1 and bneohCR2, differed from each other by 164 single-nucleotide polymorphisms (SNPs); bneohCR1 differed from the *B. neotomae* genome used as reference in the analysis (GenBank accession no. GCA_000742255) by 174 and bneohCR2 by 160 SNPs. Those data indicated that *B. neotomae* has zoonotic potential and is present in a much wider geographic area than previously reported.

Author affiliations: Université Paris-Saclay, Institute for Integrative Biology of the Cell, Gif-sur-Yvette, France (G. Vergnaud); Institut National de l'Agriculture, de l'Alimentation et de l'Environnement, Unité Mixte de Recherche Infectiologie et Santé Publique, Université de Tours, Nouzilly, France (M.S. Zygmunt, A. Cloeckaert); Animal and Plant Health Agency, Weybridge, UK (R.T. Ashford, A.M. Whatmore)

DOI: https://doi.org/10.3201/eid3001.221783

Because that finding was unexpected and has substantial implications regarding our understanding of *Brucella*, we further investigated available information regarding the neglected species *B. neotomae*. We reviewed the literature for previous studies in which *B. neotomae* strains were isolated and searched public sequence repositories for *B. neotomae* whole-genome sequence (WGS) datasets. In addition, we identified and sequenced available *B. neotomae* strains maintained since the 1960s in 2 *Brucella* strain collections, the UK Animal and Plant Health Agency (APHA) Weybridge collection and the *Brucella* Culture Collection Nouzilly (BCCN) of the Institut National de l'Agriculture, de l'Alimentation et de l'Environnement (INRAE; National Research Institute for Agriculture, Food and the Environment) in France. We report a comprehensive comparative analysis of all genome sequences we identified from databanks and the human cases from Costa Rica, to further shed light on the genetic relationships between those isolates.

The Study

We recovered 17 *B. neotomae* WGS datasets from public repositories as assemblies or raw reads (last accessed May 31, 2023): ERR1894830, GCA_000158715, GCA_000712255, GCA_000742255, GCA_900446125, SRR004305, SRR004306, SRR032598, SRR857216, SRR4038991 (all 10 strains 5K33), ERR2993140 (MLVA31), GCA_900446115, SRR4038990 (5E1169), GCA_900446105 (6D152), ERR473742 (babohCR62), ERR1845156 (bneohCR2), and ERR1845155 (bneohCR1) (Appendix Table 2, https://wwwnc.cdc.gov/EID/article/30/1/22-1783-App1.pdf). We merged 3 records (SRR004305, SRR004306, SRR032598) corresponding to the same biosample.

The *Brucella* strain collection maintained by APHA contained 5 and INRAE/BCCN, 7 *B. neotomae* strains (*6*,*7*). We recorded APHA and corresponding BCCN identifiers for each strain (Appendix Table 1). We produced and analyzed sequence data (Appendix). The 12 *B. neotomae* sequence datasets produced for

this report were SRR22273182–8 (BCCN collection corresponding to primary names 6G152, 5E1169, 5E1266, 7E1260, 6H8988, and 5G239 and 1 unknown primary name) and SRR22414766–70 (APHA collection corresponding to primary names 7E164 and 5E1266 and 3 unknown primary names). We deposited sequences in the National Center for Biotechnology BioProject database as PRJNA901374 (BCCN) and PRJNA905663 (APHA) (Appendix Table 2).

We generated a maximum parsimony tree from the 149 SNPs identified among the 27 *B. neotomae* sequence datasets, including 15 public and 12 newly sequenced WGS datasets (Figure 1; Appendix Tables 1, 2). The whole-genome SNP (wgSNP) genotype of the most recent common ancestor (MRCA) of known *B. neotomae* lineages descends into 2 groups (Figure 1): the group containing type strain 5K33 corresponds to sequence type (ST) 22 in the *Brucella* multilocus sequence typing scheme MLST21, the other to ST120 (*7*). The limited available information about the sampling site of each strain from rodents in the Great Salt Lake Desert in Utah, USA, is consistent with congruence between *B. neotomae* phylogeny and the geography of the Great Salt Lake region, but further data are needed to robustly test this association (Appendix).

We show a different representation of the wgSNP phylogenetic analysis after removal of duplicates and of 1 dataset with relatively lower coverage (Figure 2; GenBank accession no. GCA_900446105 from strain 6D152). Because we removed the WGS datasets with partial coverage, the new tree contained 205 SNPs. The distances from MRCA to tips were similar: maximum 76 SNPs (to strain 7E1260) and minimum 56 SNPs (to strain APHA#65–197). The 3 whole-genome datasets from Costa Rica, including the human isolates bneohCR1 and bneohCR2 and the isolate babohCR62 entered as *B. abortus* in the European Nucleotide

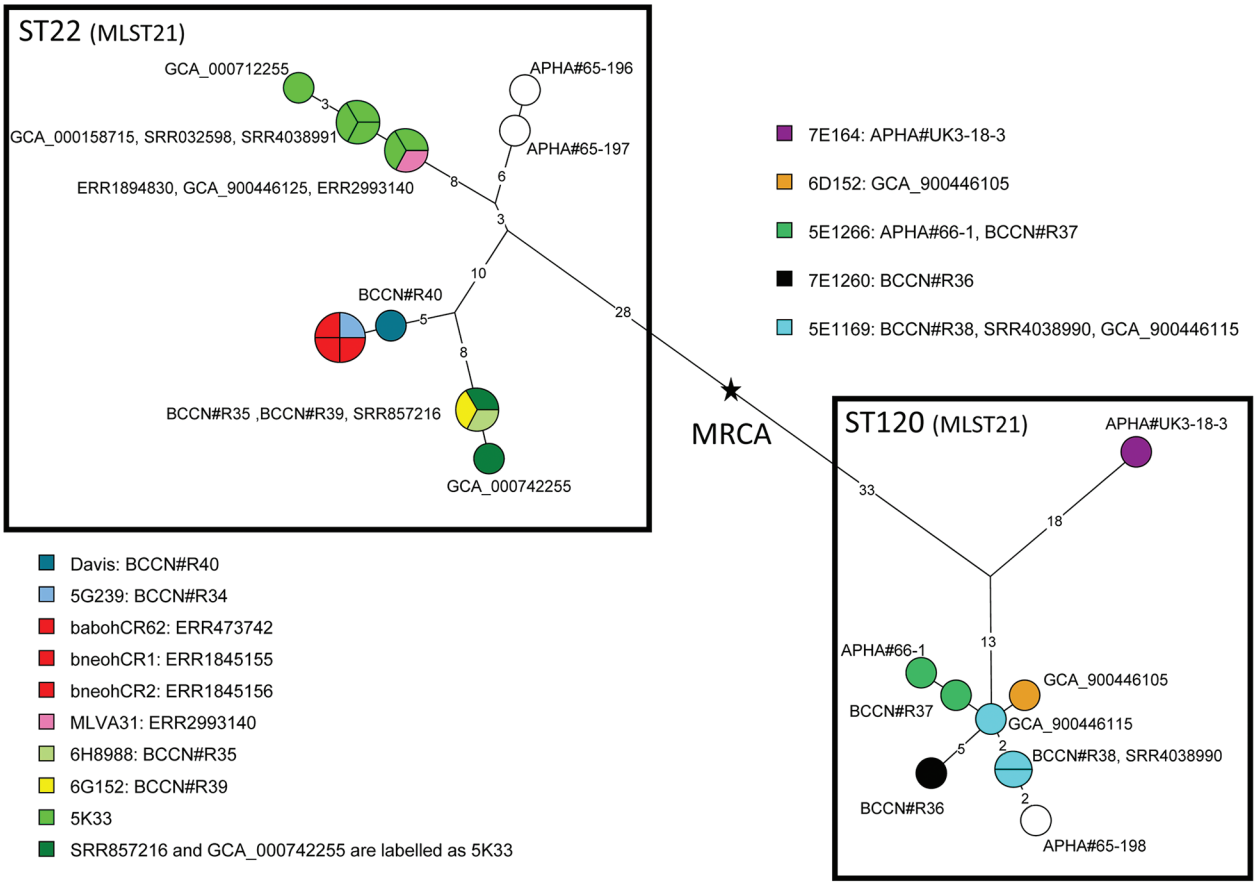


Figure 1. *Brucella neotomae* phylogeny. Maximum parsimony tree was derived from wgSNP data. We investigated 27 datasets and identified 149 SNPs; tree size is 151 substitutions (homoplasy 1.5%). Circles are colored according to primary strain identifier; red indicates the 3 datasets from Costa Rica. Circles are labeled with an accession number or collection strain identifier (*Brucella* Culture Collection Nouzilly [BCCN]) or Animal and Plant Health Agency [APHA] Weybridge collections). Branch lengths >1 substitution are indicated. Black star shows the position of the hypothetical MRCA. Box indicates the 2 MLST21 STs. MLST, multilocus sequence typing; MRCA, most recent common ancestor; SNP, single-nucleotide polymorphism; ST, sequence type; wgSNP, whole-genome single-nucleotide polymorphism.

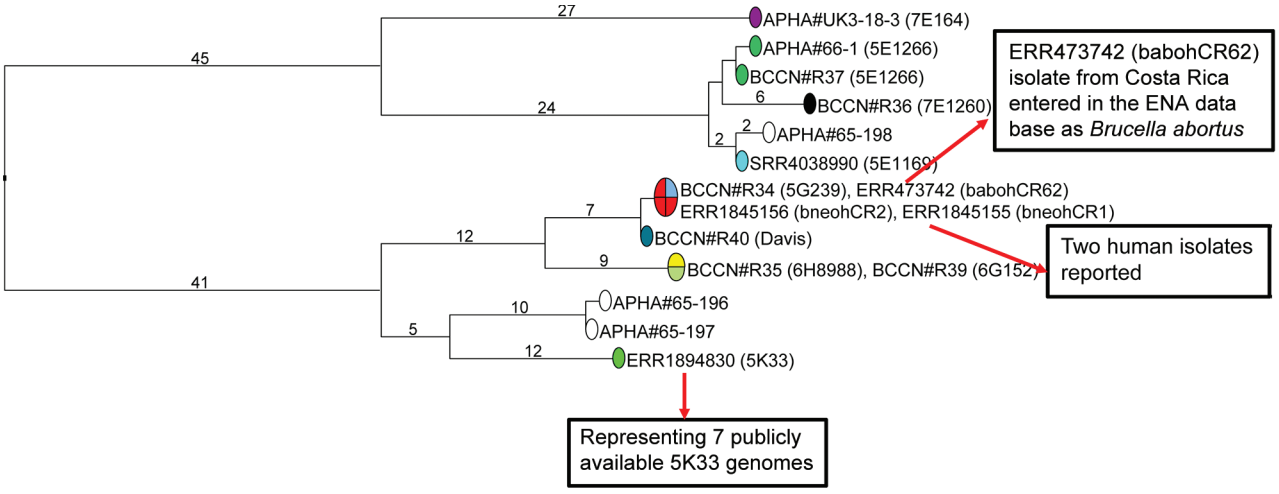


Figure 2. Rooted *Brucella neotomae* phylogeny of 16 selected datasets. Maximum parsimony tree was derived from wgSNP data; 205 SNPs in tree, tree size 207 substitutions (homoplasy 1%). Branch lengths >1 substitution are indicated. Circles are colored according to primary strain identity; red indicates the 3 datasets from Costa Rica. Circles are labeled with an accession number or collection strain identifier (*Brucella* Culture Collection Nouzilly [BCCN] or Animal and Plant Health Agency [APHA] Weybridge collections). Primary strain identifier is indicated in brackets when available. ENA, European Nucleotide Archive; SNP, single-nucleotide polymorphism; wgSNP, whole-genome SNP.

Archive database, remained identical in wgSNP genotype to strain 5G239 (BCCN#R34) in spite of the increased resolution. We still observed a coincident wgSNP genotype when we considered only these 4 strains, in sharp contrast with a report of human cases that indicated the corresponding genomes differed by 164 SNPs (*4*).

Conclusions

Our findings demonstrate that the strains isolated during 1955–1964 in the Great Salt Lake Desert in Utah display notable intraspecies genetic diversity despite being isolated from a geographically limited location, within a limited time frame, and from the same host species. In contrast, the datasets from wgSNP analysis of isolates from Costa Rica were identical despite having been isolated 4 years apart and in different areas of Costa Rica (*5*). Of note, datasets from analysis of isolates from Costa Rica were identical to data from 1 *B. neotomae* strain, 5G239, from the Great Salt Lake region. Finding an identical genotype in human cases from Costa Rica >3,000 km and >50 years apart in a different species from the Great Salt Lake discovery is remarkable in light of the diversity of strains noted in the geographically limited location in Utah and reported absence of rats of genus *Neotoma* in Costa Rica (*5*). Full understanding of the zoonotic potential of *B. neotomae* requires further exploration, including additional sampling of rodents and human cases in the US Southwest and Central America.

Work conducted at INRAE and University Paris-Saclay was supported in part by ANR (Agence Nationale de la Recherche; French National Research Agency) ASTRID-maturation project ANR-14-ASMA-0002 (Global surveillance of infectious agents: molecular typing kits and databases–MicroType). Work conducted at UK APHA was supported in part by the One Health EJP project (JRP-17–IDEMBRU: Identification of emerging Brucella species: new threats for human and animals; grant agreement no. 773830).

About the Author

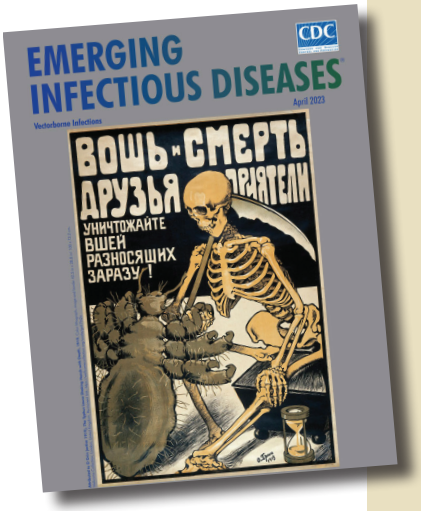
Dr. Vergnaud is a research associate at University Paris-Saclay, Gif-sur-Yvette, France. His research interests currently focus on the evolution and phylogeography of a number of dangerous bacterial pathogens.

References

1. Whatmore AM, Foster JT. Emerging diversity and ongoing expansion of the genus *Brucella.* Infect Genet Evol. 2021;92:104865. https://doi.org/10.1016/j.meegid.2021.104865
2. Whatmore AM, Perrett LL, MacMillan AP. Characterisation of the genetic diversity of *Brucella* by multilocus sequencing. BMC Microbiol. 2007;7:34. https://doi.org/10.1186/1471-2180-7-34
3. Stoenner HG, Holdenried R, Lackman D, Orsborn JS Jr. The occurrence of *Coxiella burnetii*, *Brucella*, and other pathogens among fauna of the Great Salt Lake Desert in Utah. Am J Trop Med Hyg. 1959;8:590–6. https://doi.org/10.4269/ajtmh.1959.8.590
4. Suárez-Esquivel M, Ruiz-Villalobos N, Jiménez-Rojas C, Barquero-Calvo E, Chacón-Díaz C, Víquez-Ruiz E, et al. *Brucella neotomae* infection in humans, Costa Rica. Emerg Infect Dis. 2017;23:997–1000. [Erratum in Emerg Infect Dis. 2017;23:1435.] https://doi.org/10.3201/eid2306.162018

5. Villalobos-Vindas JM, Amuy E, Barquero-Calvo E, Rojas N, Chacón-Díaz C, Chaves-Olarte E, et al. Brucellosis caused by the wood rat pathogen *Brucella neotomae*: two case reports. J Med Case Reports. 2017;11:352. https://doi.org/10.1186/s13256-017-1496-8
6. Vergnaud G, Hauck Y, Christiany D, Daoud B, Pourcel C, Jacques I, et al. Genotypic expansion within the population structure of classical *Brucella* species revealed by MLVA16 typing of 1404 *Brucella* isolates from different animal and geographic origins, 1974–2006. Front Microbiol. 2018;9:1545. https://doi.org/10.3389/fmicb.2018.01545
7. Whatmore AM, Koylass MS, Muchowski J, Edwards-Smallbone J, Gopaul KK, Perrett LL. Extended multilocus sequence analysis to describe the global population structure of the genus *Brucella*: phylogeography and relationship to biovars. Front Microbiol. 2016;7:2049. https://doi.org/10.3389/fmicb.2016.02049

Address for correspondence: Gilles Vergnaud, University Paris-Saclay, Institute for Integrative Biology of the Cell (I2BC), 1 Av de la Terrasse, Blg 24, Rm 02-335, 91190 Gif-sur-Yvette, France; email: gilles.vergnaud@universite-paris-saclay.fr



Originally published in April 2023

etymologia revisited

Haematospirillum jordania

[Hae.ma.to.spi.ril'lum jor.da'ni.ae]

For the sesquipedalian term *Haematospirillum*, *Haema* is derived from the Greek haima, meaning blood. *Spirillum* is derived from Medieval Latin in the mid-13th century Latin (*spiralis*), French in the 1550s (*spiral*), and Greek (*speira*). All suggest a winding or coil. A New Latin reference book entry in 1875 implied a little coil (Figure 1).

Isolated from human blood, Haematospirillum jordaniae was reported as a novel genus and species in 2016 by Centers for Disease Control and Prevention (CDC) scientist Ben W. Humrighouse and his laboratory team, which included Jean G. Jordan, a microbiologist (Figure 2). This gram-negative bacterium was isolated 14 times in 10 states during 2003–2012 before its identification in 2016.

H. jordaniae was previously considered an environmental bacterium with limited pathogenicity, but increasing numbers of isolates indicated a possible emerging pathogen. All cases occurred in male patients, and the pathogen showed a predilection for infecting lower leg injuries. In 2018, Hovan and Hollinger reported a case of infection in a Delaware man who, in 2016, had sepsis from a lower leg wound. The organism isolated was identified at the CDC Special Bacteriology Reference Laboratory (SBRL) in the Division of High-Consequence Pathogens and Pathology, National Center for Emerging and Zoonotic Infectious Diseases.

References

1. Hovan G, Hollinger A. Clinical isolation and identification of *Haematospirillum jordaniae*. Emerg Infect Dis. 2018;24:1955–6.
2. Humrighouse BW, Emery BD, Kelly AJ, Metcalfe MG, Mbizo J, McQuiston JR. *Haematospirillum jordaniae* gen. nov., sp. nov., isolated from human blood samples. Antonie van Leeuwenhoek. 2016;109:493–500.
3. LPSN List of prokaryotic names with standing in nomenclature. Species *Haematospirillum jordaniae* [cited 2022 May 21]. https://lpsn.dsmz.de/species/haematospirillum-jordaniae
4. Jean Jordan obituary. Published by the Atlanta Journal-Constitution on May 28, 2014 [cited 2022 Oct 17]. https://www.legacy.com/us/obituaries/atlanta/name/jean-jordan94
5. Pal E, Štrumbelj I, Kišek TC, Kolenc M, Pirš M, Rus KR, et al. *Haematospirillum jordaniae* cellulitis and bacteremia. Emerg Infect Dis. 2022;28:2116–9.
6. Persiana (1875) (Latin Edition): Heckmanns, Alexius. *Spirillum* (n.) [cited 2022 May 21]. https://www.etymoline.com/word/spirillum
7. Weyant RS, Moss CW, Weaver RE. Hollis, Jordan JG, Cook E, et al.; Centers for Disease Control and Prevention. Identification of unusual pathogenic gram-negative aerobic and facultatively anaerobic bacteria. The Orange Book, 2nd ed. Philadelphia: Lippincott Williams and Wilkins; 1996

Article DOI: https://doi.org/10.3201/eid3001.221783

Genomic Diversity and Zoonotic Potential of *Brucella neotomae*

Appendix

Materials and Methods

The APHA and BCCN strains were sequenced independently using NextSeq and MiSeq sequencing machines (Illumina, https://www.illumina.com). Genome sequence assemblies were exchanged among the participating teams and phylogenetic analyses were run in duplicate.

Whole Genome Sequencing (WGS) and Data Analysis in France

Genomic DNA was extracted from the trypticase soy agar with 0.6% yeast extract (TSAYE)-cultured strains with the DNeasy blood and tissue kit (QIAGEN, https://www.qiagen.com). Library construction and WGS were performed by GenoScreen (https://www.genoscreen.fr) on an Illumina MiSeq platform, producing 250bp long paired-end reads. Sequencing reads were assembled using SPAdes version 3.13 (*1*). The assemblies had an average of 23 contigs (range 15–33), an average N50 of 458,544 (range 252,732–653,786) and an average total assembly length of 3.30 Mb.

Assemblies were imported into BioNumerics version 8.1 (https://www.applied-maths.com). The assemblies were used to produce artificial 50 bp reads, which were mapped for SNP calling on reference genome accession number GCA_000007125 (*B. melitensis* reference strain 16M) as previously described (*2*). Maximum parsimony analysis was used for phylogenetic reconstruction from whole genome SNP (wgSNP) data. Resulting trees were rooted using the reference *B. melitensis* 16M genome accession GCA_000007125 as outgroup.

Whole Genome Sequencing (WGS) and Data Analysis in UK

APHA strains were sequenced on the NextSeq platform, using the Illumina MiSeq v2 Reagent Kit, to produce 150 bp paired-end reads. Illumina data were assembled using Unicycler version v0.4.8, which implements SPAdes version 3.13.0 for *de novo* assembly of short-read data. The quality of genome assemblies was assessed using QUAST version 5.0.2. De novo assembly of five *B. neotomae* strains from APHA Weybridge produced genomes with an average of 33 contigs (range 31–36), an average N50 of 278,142 (range 222,135–294,583) and an average total assembly length of 3.30 Mb. Assembled genomes were used for subsequent phylogenetic analysis.

Publicly available *B. neotomae* WGS data were downloaded from EBI ENA (read archives) or NCBI (genome assemblies) (last updated 31/05/2023). Sequence data produced for this study were deposited in PRJNA905663 (APHA) or PRJNA901374 (BCCN) (Supplementary Table2).

Literature Review for *B. neotomae* Isolation

A search of PubMed for "*Brucella neotomae*" OR "*B. neotomae*" in Title/Abstract retrieved 78 publications (last updated 01/11/2023). These were published between 1958 and 2021. Prior to 2017, a single investigation was quoted leading to the isolation of *B. neotomae* strains, the original investigation in the Great Salt Lake Desert, Utah, USA first reported by Stoenner and Lackman in 1957 (*3–5*). The initial report indicated that five strains were isolated (*4*) and the last available report on this effort indicated that 16 strains were recovered by 1964 (*6*). Fifteen of these strains were isolated from the desert woodrat (*Neotoma lepida)* with the 16th isolated from a flea from the same host although unfortunately no strains names are provided in this summary publication.

The sixteen *B. neotomae* strains were recovered from eight sampling sites located within an area of approximately 130x160 km^2 (*6*). Three and five sites were located on the West and East sides of the desert, respectively and an equal number of strains was recovered from each side (*6*). We also identified ten strain names quoted in different publications and following the initial naming scheme (Appendix Table 1). We were not able to find the precise correspondence between strain identifiers and sites of isolation and only partial assignments could be inferred (*6*). The initial 1957 report (*3*) quoted five strains, 5K33, 5E1169, 5E1266, 6D152 and 5G239.

The list of locations and year of sampling (*6*) indicates that three of these five were collected in Gold Hill on the West side of the Great Salt Lake desert whereas two were collected in South Cedar Mountain on the East side. Additional names were subsequently quoted in the literature (Appendix Table 1). Strains 7E164 and 6H8488 first quoted in 1958 would have been sampled in Gold Hill (West side) and Little Davis Mountain (East side) without more precision. While the authors collected more than 6000 animals from 29 different species, all *B. neotomae* strains were recovered from the 258 desert woodrat representatives. Additionally, two other *Brucella* sp. were isolated from black-tailed jackrabbits (*Lepus californicus*). This suggests that for some unknown reason *B. neotomae* appeared to be surprisingly restricted to a single host within the sampling area (*7*).

In their review published in 2014 (*8*), Olsen and Palmer mention that "about 25 cultures of *B. neotomae* have been isolated" and quote the book published in 1988 by Alton et al. (*9*). This number refers to one sentence in the introduction « *B. neotomae* was isolated from the desert wood rat (*Neotoma lepida*), a rodent that inhabits western regions of the USA. Only about 25 cultures have been isolated, none of them from domesticated animals or man». The previous edition of this book was published in 1975 and mentioned that "27 cultures have been isolated" in an otherwise identical sentence (*10*). Among the authors, Lois Jones who had been working on *B. neotomae* strains at least since 1968 was likely well informed about the number of recovered strains (*11*). In 2017, two human cases with brucellosis due to *B. neotomae* were published (*12*,*13*).

Identification of *B. neotomae* Datasets in Public WGS Repositories

We recovered 4000 core *Brucella* WGS datasets including assemblies and sequence read archives. About 750 datasets were duplicates, due to WGS data deposited as assembly and sequence read archive, or due to the independent sequencing of classical strains (including type/reference strains and vaccines strains). Whole genome SNP analysis identified a cluster of 17 closely related WGS data sets, including all datasets from the *B. neotomae* 5K33 type strain. These 17 datasets correspond to 15 biosamples (Appendix Table 2). We merged the three datasets derived from the same biosample (SAMN00102852) after checking their coincidence in terms of wgSNP. Eight biosamples are registered as corresponding to the type strain 5K33, two as strain 5E1169, and one as strain 6D152. Three are reported as originating from Costa Rica and include biosample SAMEA2266954, described as *B. abortus* strain babohCR62, in addition to

the two human isolates bneohCR01 and bneohCR02 (*12*,*13*). The *B. abortus* sample data were made public in April 2014 whereas the human isolates data were made public in June 2017 at the time of the associated publications (*12*,*13*). It is notable that one of the human Costa Rican isolates, bneohCR01, was initially described as being (mis)identified as *B. abortus* on the basis of bacteriological and biochemical tests (*12*,*13*).

Results

Whole Genome SNP Analysis of *B. neotomae* WGS Data

Six out of eight datasets assigned to the 5K33 *B. neotomae* type strain clustered together as expected (Figure 1). The one to four SNPs observed among these 5K33 representatives presumably reflect variations resulting from laboratory cultivation of the type strain or sequencing errors. One outlier 5K33 assembly (GCA_000712255) results from the assembly of IonTorrent data (which are known to have a different error profile to Illumina sequence data) (*14*) and in the absence of the raw data we could not check the quality of the detected SNPs. The last two datasets labelled as 5K33 (GCA_000742255 and SRR857216) are more than 20 SNPs away from the 5K33 group and are identical to, or one SNP away from, two coincident strains, 6G152 and 6H8988 (Figure 1). The most parsimonious explanation for this finding is that these two datasets have been incorrectly labelled as being the *B. neotomae* 5K33 type strain. Strain "MLVA31" (ERR2993140 biosample SAMEA5176147, "imported case") is most likely a 5K33 representative since it is identical to two 5K33 datasets. The three datasets corresponding to 5E1169 were separated by two SNPs. SNP distances and tree topology in Figure 1 indicate that APHA#65-198 from the APHA Weybridge collection, and missing the original strain name, is most likely strain 5E1169.

Unfortunately, because of the lack of data regarding the sampling site of each strain, it was not possible for the present time to evaluate if genotypes circulate within the whole sampling area or if on the contrary the phylogenetic tree is congruent with the geographic origin within the Great Salt Lake desert sampling area recalled in the appendix. Two among the first five strains, 5K33 and 5G239, belong to MLST21 ST22 whereas three (5E1169, 5E1266, 6D152) belong to ST120. It is tempting to speculate that the ST22 strains originated from South Cedar Mountains on the East side, whereas the ST120 strains originated from Gold Hill on the

West side of the desert. The next two strains were isolated in 1956, Little Davis Mountains or 1957, Gold Hill and correspond to 7E164 and 6H8488 according to the list provided by (*6*) and to the first appearance of these strain names in the literature (Appendix Table1) (*15*). Strain 7E164 belongs to MLST21 ST120 where it defines a distinct lineage whereas 6H8988 belongs to ST22.

References

1. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77. PubMed https://doi.org/10.1089/cmb.2012.0021

2. Vergnaud G, Hauck Y, Christiany D, Daoud B, Pourcel C, Jacques I, et al. Genotypic Expansion Within the Population Structure of Classical *Brucella* Species Revealed by MLVA16 Typing of 1404 *Brucella* Isolates From Different Animal and Geographic Origins, 1974-2006. Front Microbiol. 2018;9:1545. PubMed https://doi.org/10.3389/fmicb.2018.01545

3. Stoenner HG, Lackman DB. A new species of *Brucella* isolated from the desert wood rat, *Neotoma lepida* Thomas. Am J Vet Res. 1957;18:947–51. PubMed

4. Stoenner HG, Lackman DB. A preliminary report on a Brucella isolated from the desert wood rat, *Neotoma lepida* Thomas. J Am Vet Med Assoc. 1957;130:411–2. PubMed

5. Stoenner HG, Holdenried R, Lackman D, Orsborn JS Jr. The occurrence of Coxiella burnetil, Brucella, and other pathogens among fauna of the Great Salt Lake Desert in Utah. Am J Trop Med Hyg. 1959;8:590–6. PubMed https://doi.org/10.4269/ajtmh.1959.8.590

6. Thorpe BD, Sidwell RW, Bushman JB, Smart KL, Moyes R. Brucellosis in Wildlife and Livestock of West Central Utah. J Am Vet Med Assoc. 1965;146:225–32. PubMed

7. Meyer ME. Evolution and taxonomy in the genus *Brucella*: Brucellosis of rodents. Theriogenology. 1976;6:263–72. https://doi.org/10.1016/0093-691X(76)90019-4

8. Olsen SC, Palmer MV. Advancement of knowledge of *Brucella* over the past 50 years. Vet Pathol. 2014;51:1076–89. PubMed https://doi.org/10.1177/0300985814540545

9. Alton GG, Jones LM, Angus RD, Verger JM. Techniques for the brucellosis laboratory. Paris, France: Institut National de la Recherche Agronomique (INRA); 1988.

10. Alton GG, Jones LM, Pietz DE. Laboratory techniques in brucellosis. 2nd ed. Geneva: World Health Organization; 1975.

11. Jones LM, Merz GS, Wilson JB. Phage typing reactions on *Brucella* species. Appl Microbiol. 1968;16:1179–90. PubMed https://doi.org/10.1128/am.16.8.1179-1190.1968

12. Suárez-Esquivel M, Ruiz-Villalobos N, Jiménez-Rojas C, Barquero-Calvo E, Chacón-Díaz C, Víquez-Ruiz E, et al. *Brucella neotomae* Infection in Humans, Costa Rica. Emerg Infect Dis. 2017;23:997–1000. PubMed https://doi.org/10.3201/eid2306.162018

13. Villalobos-Vindas JM, Amuy E, Barquero-Calvo E, Rojas N, Chacón-Díaz C, Chaves-Olarte E, et al. Brucellosis caused by the wood rat pathogen *Brucella neotomae*: two case reports. J Med Case Reports. 2017;11:352. PubMed https://doi.org/10.1186/s13256-017-1496-8

14. Salipante SJ, Kawashima T, Rosenthal C, Hoogestraat DR, Cummings LA, Sengupta DJ, et al. Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. Appl Environ Microbiol. 2014;80:7583–91. PubMed https://doi.org/10.1128/AEM.02206-14

15. Cameron HS, Meyer ME. Metabolic studies on *Brucella neotomae* (Stoenner and Lackman). J Bacteriol. 1958;76:546–8. PubMed https://doi.org/10.1128/jb.76.5.546-548.1958

16. Parnas J, Zalichta S, Sidor-Wójtowicz A. The taxonomic properties of *Brucellae* isolated from *Neotoma lepida* Thomas (*Brucella neotomae* Stoenner). Zentralbl Veterinärmed B. 1967;14:634–45. PubMed https://doi.org/10.1111/j.1439-0450.1967.tb00266.x

17. Meyer ME, Cameron HS. Comparative metabolism of species and types of organisms within the genus *Brucella.* J Bacteriol. 1959;78:130–6. PubMed https://doi.org/10.1128/jb.78.1.130-136.1959

18. Hall WH, Manion RE. In vitro susceptibility of *Brucella* to various antibiotics. Appl Microbiol. 1970;20:600–4. PubMed https://doi.org/10.1128/am.20.4.600-604.1970

19. Stoenner HG. Observations on the Behavior of *Brucella neotomae* in Laboratory Animals. Am J Vet Res. 1965;26:347–52.

20. Freeman BA, Baughn RE, McGhee JR. Some physical, chemical, and taxonomic features of the soluble antigens of the *Brucellae.* J Infect Dis. 1970;121:522–7. PubMed https://doi.org/10.1093/infdis/121.5.522

21. Beal GA, Lewis RE, McCULLOUGH NB, Claflin RM. Experimental infection of swine with *Brucella neotomae.* Am J Vet Res. 1959;20:872–5. PubMed

22. Stoenner HG. The behavior of *Brucella neotomae* and *Brucella suis* in reciprocal superinfection experiments in mice and guinea pigs. Am J Vet Res. 1963;24:376–80. PubMed

Appendix Table 1. List of *B. neotomae* strains identified in the literature and additional isolates identified in historical strain collections

Strain ID	Collection aliases	Quoting references	WGS data available	Comments
5K33	NCTC 10084 ATCC 23459	(*3*,*7*,*11*,*15*–*19*)	SRR032598[$] SRR004305[$], SRR004306[$] SRR857216* SRR4038991 GCA_000158715 GCA_000742255* GCA_000712255 GCA_900446125 ERR1894830 ERR2993140[£]	[$] the three datasets correspond to the same biosample SAMN00102852, share the same wgSNP genotype and were subsequently merged under the name * GCA_000742255 and SRR857216 do not cluster with the other 5K33 datasets "SRR032598" in Figure 1 [£] MLVA31, biosample SAMEA5176147 incorrectly labelled "imported case, isolated in Germany" (Dr Enrico Georgi, personal communication)
5E1169	NCTC 10070, BCCN#R38	(*3*,*7*,*11*,*15*,*16*,*19*)	SRR4038990 GCA_900446115 This report	BCCN strain transmitted by Dr L. Jones
5E1266	BCCN#R37 APHA#66-1	(*3*,*11*,*15*–*17*)	This report	Mistranscribed as SE1266 in (*16*). Received in APHA Weybridge collection in 1966.
6D152	NCTC 10071	(*3*,*15*,*16*,*19*)	GCA_900446105	
5G239	BCCN#R34	(*3*,*7*,*11*,*15*)	This report	Presumed mistranscribed as 56-239 in (*20*)
7E164	NCTC 10072, APHA#UK3-18-3	(*7*,*11*,*15*,*18*,*19*,*21*,*22*)	This report	Presumed mistranscribed as 7E1164 in (*11*)
6H8488	BCCN#R35	(*7*,*11*,*15*)	This report	Presumed mistranscribed as 6H8988 in BCCN
6G152	BCCN#R39	(*7*,*11*)	This report	Presumed mistranscribed as 6G150 in (*11*) and 66-152 in (*20*). Transmitted by Dr L. Jones
7E1260	BCCN#R36	(*7*)	This report	
6D239		(*16*)		
Unknown	1963	(*7*)		A single strain was isolated in 1963,on the East side (*6*)
Unknown	239, 3n, 4n	(*16*)		
Unknown	BCCN#R40 (Davis)		This report	Transmitted by Dr L. Jones. Dr M. Meyer was working at University of California, Davis. Potential source? Possibly "1963" the only strain quoted by Meyer and missing in the BCCN and APHA Weybridge collections
Unknown	APHA#65-196	PubMLST	This report	Received in APHA Weybridge collection in 1965 from Dr Thorpe, Utah, USA. No additional identifier recorded.
Unknown	APHA#65-197	PubMLST	This report	Received in APHA Weybridge collection in 1965 from Dr Thorpe, Utah, USA. No additional identifier recorded
Unknown	APHA#65-198		This report	Received in APHA Weybridge collection in 1965 from Dr Thorpe, Utah, USA. No additional identifier recorded.

The original name of 8 among the 12 strains was recorded and corresponded to names identified in the literature except for 1 presumed transcription error.

Appendix Table 2 WGS datasets derived from *B. neotomae* strains

Key	BioProject	BioSample	Indicated strain	Country
ERR1894830	PRJEB19503	SAMEA103935299	5K33	USA
GCA_900446125	PRJEB6403	SAMEA104210778	NCTC10084 (5K33)	USA
GCA_900446115	PRJEB6403	SAMEA104318192	NCTC10070 (5E1169)	USA
GCA_900446105	PRJEB6403	SAMEA104318193	NCTC10071 (6D152)	USA
ERR473742	PRJEB4782	SAMEA2266954	babohCR62	Costa Rica
ERR1845156	PRJEB19503	SAMEA94360168	bneohCR02†	Costa Rica
ERR1845155	PRJEB19503	SAMEA94360918	bneohCR01†	Costa Rica
ERR2993140	PRJEB30030	SAMEA5176147	MLVA31	Germany
SRR004305	PRJNA33567	SAMN00102852	5K33	USA
SRR004306	PRJNA33567	SAMN00102852	5K33	USA
SRR032598	PRJNA33567	SAMN00102852	5K33	USA
SRR857216	PRJNA194124	SAMN01990992	5K33	USA
GCA_000712255	PRJNA230241	SAMN02427357	5K33	USA
GCA_000158715	PRJNA33567	SAMN02595289	5K33	USA
GCA_000742255	PRJNA243897	SAMN02768006	5K33	USA
SRR4038990	PRJNA251693	SAMN05417903	5E-1169	USA
SRR4038991	PRJNA251693	SAMN05417904	5K33	USA
SRR22273188*	PRJNA901374	SAMN31711931	5G-239	USA
SRR22273187*	PRJNA901374	SAMN31711932	6H-8988	USA
SRR22273186*	PRJNA901374	SAMN31711933	7E-1260	USA
SRR22273185*	PRJNA901374	SAMN31711934	5E-1266	USA
SRR22273184*	PRJNA901374	SAMN31711935	5E-1169	USA
SRR22273183*	PRJNA901374	SAMN31711936	6G-152	USA
SRR22273182*	PRJNA901374	SAMN31711937	Davis	USA
SRR22414766*	PRJNA905663	SAMN31880432	NCTC10072 (UK3/18-3) (7E164)	USA
SRR22414767*	PRJNA905663	SAMN31880431	66/1 (5E1266)	USA
SRR22414768*	PRJNA905663	SAMN31880429	65/197	USA
SRR22414769*	PRJNA905663	SAMN31880430	65/198	USA
SRR22414770*	PRJNA905663	SAMN31880428	65/196	USA

*This report

†Called bneohCR1 and bneohCR2 in reference (*12*).