



OPEN Investigation of genomic island 2 deciphers the evolution of a genus *Brucella*

Menachem Banai

HGT acquisition of genomic island 2 (GI-2) into the *Brucella* genome endowed these organisms with two transglycosylases which function in the final steps of the O-polysaccharide (OPS) polymerization. Here, the investigation of the brucellaphage Pr genome revealed sequence similarities to GI-2 in two DNA regions, the *Chelativorans* sp. BNC1 and the *Ochrobactrum anthropi* ATCC 49188 DNAs, respectively. This led the investigation of the published GI-2 elements, using the phage *attB* and *attP* sequences as hallmarks of the integration site. The study identified a *Mesorhizobium loti* integrative, conjugative element (ICE) into which GI-2 integrated. A similar ICE, which nevertheless lacks GI-2, and shows characteristic chromosomal architectural differences between the taxa, was found in the genomes of the *Ochrobactrum* sister clade and the alphaproteobacterial *Chelativorans* sp. BNC1 strains. This proves that *Brucella* and *Ochrobactrum* diverged at the genus level while being in the state of a monopartite genome bearing organism.

Keywords *Brucella*, *Ochrobactrum*, Genus, Integrative conjugative element, Chromosome, tRNA

The problem of brucellosis has first been documented as an undulant, Mediterranean sickness in the Maltese population (Malta Fever). Sir David Bruce was first in linking the death of four British Royal Army soldiers from the deployed forces in the island to a coccus-shaped *Micrococcus melitensis* bacterium¹. Themistocles Zammit, a member of the Mediterranean Fever Commission (MFC), subsequently established the zoonotic nature of these bacteria following their isolation from the goat's milk². Isolation of '*Bacillus abortus*' from the placenta of a cow with contagious abortion³ and then, isolation of a somewhat similar bacterium, 'porcine abortus strain' from an aborted fetus of a sow⁴ pioneered the work of Evans⁵ who first indicated that these bacteria, which were isolated from different animal sources, are in fact similar organisms. Accordingly, Meyer and Shaw⁶ designated a new taxonomical rank, genus *Brucella*, which included *Brucella abortus* and *B. melitensis*, and *B. suis* was later recognized as a separate species⁷.

Isolation of *B. ovis* which infects rams and causes abortions in pregnant ewes and *B. canis* which causes abortions in pregnant bitches, both bearing a stable rough form, and a smooth *B. neotomae* which infects wood rats, confirmed the iconic association of these bacteria to a specific natural host, in which they cause abortion at third trimester⁸. The growing numbers of species with advert characteristics led Corbel and Morgan propose minimal standards for inclusion of new isolates to the genus⁹. However, the finding that the DNA/DNA hybridization values between the genomes of the six known species at the time are below the limits which justify the nomination of a species their inclusion in a single *B. melitensis* monospecific single taxon was proposed¹⁰. In response, the Sub-Committee on the taxonomy of *Brucella* rebutted this concept, partly due to overlooking the significance of the genetic diversion that these organisms bear¹¹.

Identification of novel isolates from rodents in North Queensland, Australia¹², an atypical strain of *B. suis* bv. 5¹³, *B. abortus* bv. 7¹⁴, marine mammal strains from cetaceans and pinnipeds which resemble "classical *Brucella*"¹⁵ as well as *Brucella vulpis* from mandibular lymph nodes of red foxes¹⁶ and *Brucella papionis* from baboons¹⁷ further highlighted the fact that *Brucella* species differ by their genomes^{11,15}. Then after, isolation of a *B. microti* soil isolate which was later confirmed to be a multi-host pathogen, with faster growth characteristics than "classical *Brucella*" and phenotypically similar to *Ochrobactrum*¹⁸ as well as finding cross infective strains between wild animals and humans^{19–21}, and amphibian strains in addition²², indicated the existence of a new group of atypical strains which are phylogenetically closer to *Ochrobactrum*²³. *B. inopinata* BO1 which was isolated from an artificial breast implant²⁴, and its animal source remains unknown until to date, presents a type strain for the group, with the genome including unique molecular patterns such as presence of a temperate phage²⁵.

Department of Bacteriology, Kimron Veterinary Institute, Bet Dagan, Israel. email: menachemba48@gmail.com

The development of a breadth of high throughput molecular techniques enabled resolution between species and strains^{26–28}. Nevertheless, results with these methods disagree with the classical bacteriological typing of biovars within species²⁹, and from a plain cladistic approach, molecular data support inclusion of *Ochrobactrum* in genus *Brucella*³⁰. This navigation between criteria which define a genus³¹ exemplified a confusion how the scientific nomenclature affected the reporting system according to which clinical laboratories notify the occurrence of *Ochrobactrum* and *Brucella* human cases³².

Brucella lipopolysaccharide (LPS) composes of an inner lipid A of very long chain fatty acids (VLCFA)^{33–35}, and a branched oligosaccharide LPS core³⁶, which Cardoso et al. and Roop et al. reported hinder the cells from the host's immune response^{37,38}. Outside, *Brucella* O-polysaccharide (OPS) establishes a major antigen which is instrumental in serological diagnosis of *Brucella* infection, and our studies showed the antigen activates development of antibodies which are functional in cell killing by the complement mediated pathway³⁹. This antigen is structured as a homopolymer of α -(1,2) glycosidic linkages between perosamine units which at the non-reducing end includes one or more tetrameric units in which the third bond is α -(1,3)⁴⁰.

The functions of perosamine synthesis and OPS polymerization have been acquired via horizontal gene transfer (HGT) of *wbk* and genomic island-2 (GI-2), respectively. *wbk* which includes *manA*_{OAg}, *manB*_{OAg} (*pmm*), *manC*_{OAg}, *gmd* and *per*, as well as *wbkA*, *wbkB* and *wbkC*^{41–43} participates in the synthesis of perosamine. BMEI1326, a glycosyltransferase Family 25 protein (first annotated *wa*** and later *wadA*⁴⁴), is yet an additional gene which resides out of *wbk* and plays an unresolved role in synthesis of the inner core LPS⁴³. GI-2, which was first documented by Chain et al. as pathogenicity island PAIB1 in *B. abortus* 2308⁴⁵, includes two transglycosylases encoding genes, *wboA* and *wboB*⁴⁶, which participate in the final steps of OPS polymerization, as well as BAB1_0989 and BAB1_0990 which encode two outer membrane proteins (OMPs) important in suppressing the host's cytokine response³⁸. Significantly, *Ochrobactrum* is missing *wbk* but individual gene orthologs of this element are distributed along its chromosome I (ChrI), suggesting acquirement of the element in early evolution and gene reduction throughout speciation later⁴⁷.

Here, a mobile integrative and conjugative element (ICE) in the *Mesorhizobium loti* genome is shown to be horizontally transferred into the *Ochrobactrum* and *Brucella* genomes, linking their evolution to an archaeal monopartite genome bearing organism. It is inferred that the archetype *Brucella* evolved from this organism and the twin sister *Ochrobactrum* evolved as its daughter cell. Evidence shows that the GI-2 integration to the *Brucella* ICE occurred only after this event, confirming the genus position of the two clades in the phylogenetic tree.

Results

The paste and cut mechanism of gene development

The brucellaphage Pr genome is a circularly permuted DNA of 38,253 bp which is composed of two opposing DNA blocks, one originating in a *Listonella* phage ϕ HSIC DNA and encoding DNA metabolizing genes and the second originating in a *Chelativorans* sp. BNC1 cryptic prophage and encoding morphogenesis and host lysis genes⁴⁸. Figure 1 and Table 1 show, in addition, two Phage Pr pillar sequences which anchor GI-2 sequences to *Ochrobactrum* and *Chelativorans* sp. BNC1 precursor cells from which contemporary *Ochrobactrum* and *Brucella* evolved. The first sequence involves an intergenic sequence between ORF 11 and ORF 12 in the *Chelativorans* sp. BNC1 cryptic phage DNA region which has sequence similarity to an intergenic sequence between BAB1_0990 and BAB1_0991 of PAIB1. The second sequence overlaps the complete sequence of ORF 33 and a partial sequence of ORF 34, which portray sequence similarity to an intergenic sequence between Oant_0203 and Oant_0204 in the *O. anthropi* ATCC 49188 genome. However, as can be seen in Fig. 1 and Table 1, 33 out of 36 nts. of this sequence also have similarity to BAB1_0984 in the *B. abortus* strain 2308 PAIB1. There are, in addition, seven other short sequences similarities to Oant_0204 and Oant_0207, which link the phage genome to the *O. anthropi* ATCC 49188 genome at this site.

Evolution of GI-2 genes in core *Brucella* species

The *B. ovis* scar sequence was used to identify the borderlines where GI-2 elements integrated into the studied genomes. This identified the specific genes which reside within GI-2, allowing to compare the sequence of each gene to the *B. abortus* strain 2308 genome, BLASTn querying *Brucella abortus* 2308 (taxid:359391). Table S1 depicts the chromosomes in where GI-2 is found in the *Brucella* genomes. As can be seen, most *Brucella* strains have their elements included in ChrI, but others have the element in their ChrII. Sequence comparisons between the genes of an annotated genome and the reference *B. abortus* strain 2308 show that the GI-2 elements differ by their gene content as well as their sizes, and may exist in the form of scar. For example, BAB1_0982 only exists as an intergenic non-coding sequence between BMEI24 (*tRNA*^{Gly}) and BMEI1013 (hypothetical membrane protein) in *B. melitensis* strain 16 M; BAB1_0998 upstream *wboA* (hypothetical protein, a functional protein in OPS synthesis) is longer than its BMEI0999 hypothetical protein homolog by 93 nts.; BAB1_0997 (conserved hypothetical protein, a phage related protein) is 438 nts. long and its BMEI1000 homolog is 738 nts. Further, *B. suis* bv. 3 DK67_1871 (outer membrane autotransporter barrel domain protein) and some extra nucleotides comprise in *B. abortus* 2308 two OMPs, BAB1_0989 and BAB1_0990; DK67_1878 (phage tail family protein) is 339 nts. longer at its 5'-end than BAB1_0997 (conserved hypothetical protein), and in comparison, BMEI1000 in *B. melitensis* bv. 1 str. 16 M is longer than BAB1_0997 by 309 nts.; DK67_1876 (hypothetical protein) includes *B. abortus* 2308 IS711 (BAB1_0996) as an intragenic sequence, and a novel DK67_1887 transposase DDE domain protein is missing in the other *Brucella*. In *B. neotomae* 5K33 (Contig188 DK64_1987, <https://www.ncbi.nlm.nih.gov/nuccore/JMSC01000003.1>) GI-2 seems to be similar to that of *B. suis* bv. 3, strain 686, but its two tRNA-guanine transglycosylase family protein genes (DK64_1989 and DK64_1988) establish a composite sequence in the former (DK67_1875, /product = "tRNA-guanine transglycosylase family protein"), (data not shown).

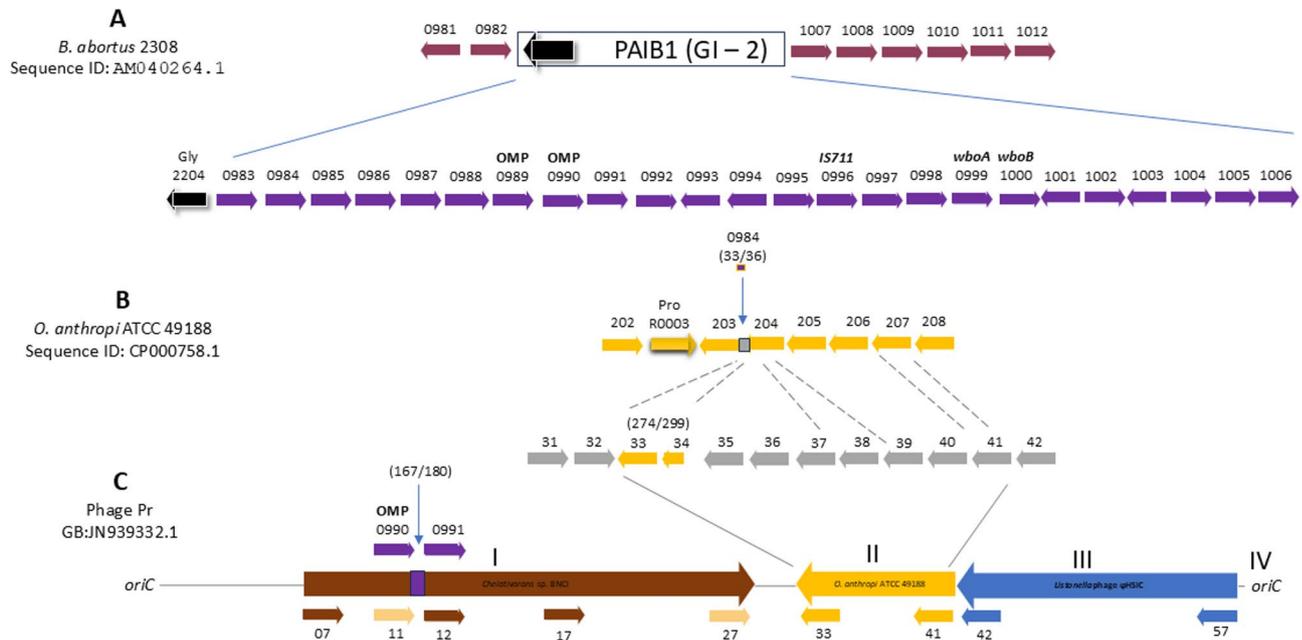


Fig. 1. The genetic maps of the studied DNA regions and the corresponding sequence similarities (The maps are not to scale, sequences are aligned): (A) The *B. abortus* strain 2308 PAIB1 (GI-2 in other *Brucella*) and the flanking genes of this element (Top) and the specific genetic map of PAIB1 (Bottom). The black shadowed arrow depicts BAB1_2204 (*tRNA^{Gly}*, *attB*). (B) The specific DNA unit in chromosome 1 of *O. anthropi* ATCC 49188. (C) The Phage Pr ORFs which have sequence similarities to the indicated *O. anthropi* ATCC 49188 genes (top, broken lines point to the specific chromosomal sequences where similarities are found, see Table 1 for the specific parameters) and the genetic map of Phage Pr (Bottom). I – The Phage Pr DNA which corresponds to the *Chelativorans* sp. BNC1 prophage (The specific Phage Pr ORFs which have partial sequence homology to *Chelativorans* sp. BNC1 are colored Brown, and those which have amino acid similarity are colored Naples yellow reddish). II – The Phage Pr DNA which has sequence similarities to the *O. anthropi* ATCC49188 chromosome 1 (Ochre, arrows below indicate the Phage Pr ORFs which mark the borderlines, *ibid*, see Table 1 for parameters); III – The Phage Pr DNA which corresponds to the *Listonella* phage ϕ HSIC genome (light Blue arrows below indicate the Phage Pr ORFs which mark the borderlines). IV-*oriC* stands for the replication origin (genome is circulated⁴⁸). Arrows stand for the specific genes and the heads show the direction of the gene transcription. Numbers above the arrows depict the specific locus tags per a strain (*B. abortus* 2308, and *O. anthropi* ATCC 49188) and ORFs (Phage Pr), respectively. The genes are colored per an organism, e.g., purple is assigned to PAIB1, ochre is assigned to *O. anthropi* ATCC 49188, and grey is assigned to Phage Pr. The shadowed arrows depict the specific tRNA genes. The vertical arrows point to the specific DNA region in which sequence similarity is shown and the word size is given in parenthesis.

Evolution of GI-2 genes in unclassified *Brucella* species

The GI-2 gene content of atypical and unclassified *Brucella* strains diverge significantly from the one found in core *Brucella*. As can be seen in Table S1 GI-2 exists in several forms which are arbitrarily coned modified elements e.g. as depicted in *Brucella* sp. 2002734562; *Brucella* sp. 2594; and *Brucella* sp. 1315. Alternatively the modified GI-2 may be novel as shown in *Brucella* sp. 2280; *Brucella* sp. 09RB8471 and *Brucella* sp. 10RB9215 isolate BR10RB9215. In comparison *Brucella* sp. NM4 possesses *tRNA^{Gly}* (*attB*) in chromosome 2 but lacks GI-2 nor does it have scar (*attP*) and therefore resembling *Ochrobactrum*. Yet the DNA architecture which corresponds to the tRNA genes upstream and nitrogen metabolizing genes downstream resembles that of *Brucella* (see below). Moreover as shown in Figs. 2 and 3 all strains share a conserved DNA motif (see Figures S1 and S2 for the specific sequences of this DNA unit in the *B. abortus* strain 2308 and *O. anthropi* ATCC 49188 chromosomes) which is also present in this strain but the sequence lacks similarity to the other *Brucella*-all but the last two genes: dTMP kinase (/locus_tag="QLQ09_13860") and DNA polymerase III subunit delta' (/locus_tag="QLQ09_13865"). Strikingly *Brucella* sp. MAB-22 and *Brucella* sp. JSBI001 lack the GI-2 attributes at the expected *tRNA^{Gly}* (*attB*) site and the upstream region comprises a similar architecture as that of the *Ochrobactrum* species (Fig. 3 see below) suggesting that they are misplaced in this tree.

Scar

B. ovis 63/290 ATCC 25840 is a naturally stable rough *Brucella* strain and *B. abortus* 2308 *B. suis* bv.3 strain 686 and *B. melitensis* bv. 1 strain 16 M among others are smooth LPS types. Transformation from smooth to rough LPS has been explained due to a precise excision of GI-2 from the *B. ovis* chromosome and development of scar at this site⁴⁹. Figure 2 depicts the genetic maps of the aligned GI-2 elements and flanking genes of the core and the atypical *Brucella* as derived from BLASTn analyses of the *B. ovis* scar sequence⁴⁶. *O. anthropi* ATCC 49188

Phage Pr nt position	Corresponding ORF	<i>O. anthropi</i> nt position	Corresponding CDS	<i>B. abortus</i> 2308 nt position	Corresponding CDS
24062..24360	33, 34 (ATG)	226080..226378 (+/+); 274/299 Expected 6e-118	Between Oant_0203 and Oant_0204	955138..955174 (+/+); 33/36 (3 mismatches at center)	BAB1_0984
25335..25392	37(ATG), 38(Ter)	226966..227023 (+/+); 56/58 Expected 1e-18	Oant_204		
25,491..25536	None	227006..227051 (+/+); 46/46 Expected 9e-15	Oant_204		
25616..25667	39(Ter)	227357..227408 (+/+); 46/52 Expected 7e-10	Oant_204		
25940..25981	41(Ter)	227,696..227737 (+/+); 41/42 Expected 6e-11	Oant_207		
26110..26149	41(ATG)	227809..227848 (+/+); 37/40 Expected 3e-07	Oant_207		
26190..26233	41(core)	228,186..228229 (+/+); 41/44 Expected 2e-09	Oant_207		
26765..26813	41(ATG)	228768..228817 (+/+); 46/50 1Gap Expected 2e-10	Oant_207		
5923..6102	Between ORF_11 and ORF_12			958473..958651 (+/+);167/180 Gaps 1 Expected 9e-68	Between BAB1_0990 and BAB1_0991

Table 1. Sequence similarities between the genome of Phage Pr, *B. abortus* 2308 PAIB1 and *O. anthropi* ATCC 49,188 chromosome 1. ATG and Ter stand for start and termination sequences and core for central sequence. Pluses in parentheses indicate sequence polarity. Numbers indicate sequence identity.

is shown at top as a representative strain of the taxonomically closest clade which lacks GI-2 and neither has scar (Table S2). As can be seen BAB1_1006 which includes *attP* marks the end of PAIB1 in *B. abortus* 2308 and BAB1_1007 lies outside the element. In contrast BMEI0993 which marks the end of GI-2 in *B. melitensis* bv. 1 strain 16 M establishes a composite sequence of the BAB1_1006 and BAB1_1007 CDS and therefore *attP* is intragenic in this ORF. Such an amorphic *attP* sequence is best illustrated in *B. suis* bv. 3 strain 686. In this organism DK67_1885 is the last gene within GI -2 and the sequence overlaps BAB1_1004 and an intergenic sequence between BAB1_1004 and BAB1_1005. However GI-2 terminates with a sequence which links this gene to DK67_1886 *arsC* out of GI-2 and corresponds to an intragenic sequence between BAB1_1006 and BAB1_1007 which includes *attP*.

Data shown in Table S1 identify conserved scars in ChrI of *Brucella* sp. 09RB8910; *Brucella* sp. 2716; *Brucella* sp. 458; and *Brucella* sp. 6810. As can be seen in Fig. 2 and Table S1, unexpectedly, a conserved scar is also depicted in ChrII of *B. inopinata* BO2. In contrast, as can be seen in Fig. 2, scar in ChrI of *B. inopinata* BO 1 is composed of a truncated *attP* (GCGG) sequence adjunct to the *tRNA^{Gly}* CDS and 39 nts. apart is a reduced sequence of BAB1_1006 which still includes an intact *attP*. Finally, as shown in Fig. 2 and Table S1, GI-2 in *B. inopinata* BO3 is novel, lacking iconic cognate phage integrase as well as phage *attP*.

Gene synteny at the *tRNA^{Gly}* chromosomal region

Gene synteny at the *tRNA^{Gly}* up—and downstream flanking genes is depicted in Fig. 2. As can be seen *O. anthropi* ATCC 49188 and *Brucella* share a conserved gene cluster immediately downstream *tRNA^{Gly}* (which in *Brucella* is part of GI -2) however their pair genes do not have significant sequence similarity (data not shown). Upstream *O. anthropi* ATCC 49188 differs from *Brucella* the former having a conserved gene cluster associated to nitrogen metabolizing genes which in core and atypical *Brucella* is replaced by a characteristic *tRNA^{Arg}* and *tRNA^{Val}* gene region. Data shown in Fig. 3 further confirm that the missing genes in one are present in the opposite side in the other respectively.

A survey of the genes along the studied chromosomes of all the *Ochrobactrum* strains then confirmed that they share the same gene synteny (data not shown). Among *Brucella*, exceptionals include *Brucella* sp. 458 which depicts a similar genetic map as *B. suis* bv. 3, strain 686, but GI-2 is excised (conserved scar, ChrI), and *B. inopinata* BO2 (conserved scar, ChrII) in which the nitrogen metabolizing gene cassette is being located in ChrI (data not shown). Other exceptionals include *Brucella* sp. 09RB8910 (ChrI) and *Brucella* sp. 09RB8471 (ChrII) in which the nitrogen metabolizing genes are positioned *Ochrobactrum*—like upstream *tRNA^{Gly}*, and *Brucella* sp. 10RB9215 isolate BR10RB9215 which has the genes ordered as in *B. suis* bv.3, strain 686 chromosome (data not shown).

BLASTn analysis of the *B. abortus* strain 2308 gene sequences cross checked against *Brucella* species and *O. anthropi* ATCC 49188 and vice versa revealed that the conserved genes are homologous between self but not

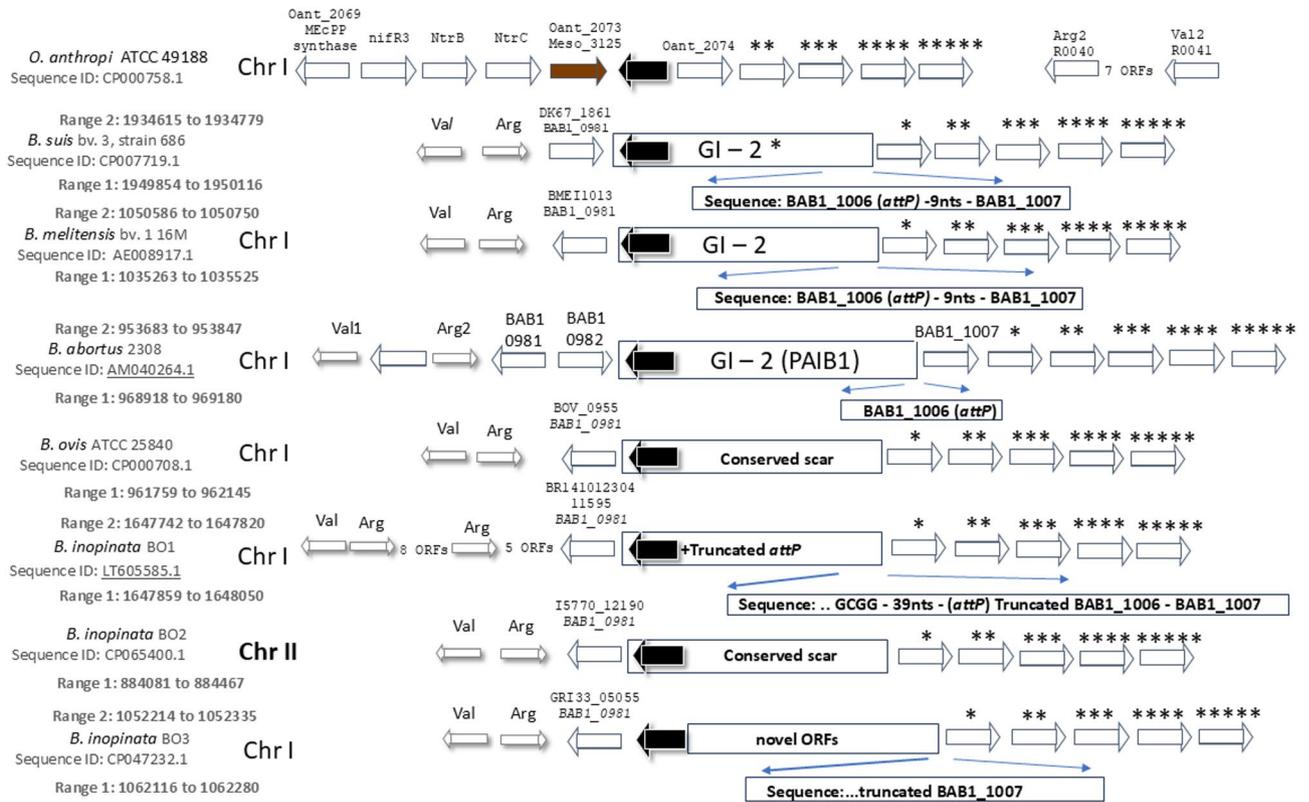


Fig. 2. The genetic maps of the DNA region at the *tRNA^{Gly}* locus (The maps are not to scale, sequences are aligned). Top-*O. anthropi* ATCC 49188; Second-*B. suis* bv. 3, strain 686; Third-*B. melitensis* bv. 1, strain 16M; Fourth-*B. abortus* 2308; Fifth-*B. ovis* ATCC 25840; Sixth-*B. inopinata* BO1, strain 141012304; Seventh-*B. inopinata* BO2; Last-*B. inopinata* BO3. Stars stand for the annotation of the *B. abortus* 2308 gene functions: locus_tag = “BAB1_1008” (= “Arsenate reductase: Arsenate reductase and related”), locus_tag = “BAB1_1009” (= “Rare lipoprotein A”), locus_tag = “BAB1_1010” (= “D-alanyl-D-alanine carboxypeptidase 1, S11 family”), locus_tag = “BAB1_1011” (= “Thymidylate kinase: ATP/GTP-binding site motif A (P-loop)”) and locus_tag = “BAB1_1012” (= “Replication factor C conserved domain”). *O. anthropi* ATCC 49188 has all but the first gene which is substituted for / locus_tag = “Oant_2074”; /product = “diacylglycerol kinase catalytic region” (See Figure S2 for the sequence of this DNA unit). Black shadowed arrows stand for the specific *tRNA^{Gly}*, free or included in the GI-2 element (rectangular box). The specific sequence which terminates GI-2 and includes *attP* is indicated below each GI-2, in a second rectangular box (see Table S1). The chromosome which bears the queried DNA is indicated by its number. Sequence ID and ranges of sequence similarities (e.g., finding *attB* and *attP*) are indicated per each strain. The arrows stand for the specific genes and the heads show the direction of the gene transcription. The shadowed arrows stand for the specific tRNA CDS.

between the heterologous group except for the *ispD* and *ntnC* external genes which share sequence similarities also to the heterologous ones (denoted by a star in Fig. 3). As the gene functionalities are similar differences between sequences indicate development of orthologous genes at this region.

Introducing *Brucella* sp. JSB1001

A consensus has been agreed upon the unique position of *B. suis* bv. 3 strain 686 among *Brucella* as a monopartite genome bearing organism^{50,51}. As can be seen in Table S1 the analysis identified that *Brucella* sp. JSB1001 has a single chromosome. BLASTn analysis of *B. ovis* scar against unclassified *Brucella* (taxid:2632610) identified strain JSB1001 however lacking GI-2 or scar. As can be seen in Fig. 3 chromosomal walking along the genes at this site then showed that the DNA overlaps that of *O. anthropi* ATCC 49188. This result positions *Brucella* sp. JSB1001 as the first identified *O. anthropi* monopartite genome organism.

How is *Chelativorans* sp. BNC1 involved?

As shown in Fig. 2 /locus_tag = “Oant_2073”/product = “hypothetical protein” /note = “KEGG: mes:Meso_3125 hypothetical protein” is the first gene upstream *tRNA^{Gly}* in the *O. anthropi* ATCC 49188 DNA. An intuitive search found 40 more randomly distributed Meso – associated genes in *O. anthropi* ATCC 49188 chromosome 1 (ChrI) and 61 other Meso – associated genes in chromosome 2 (ChrII). Worth mentioning are Oant_0390 (415547..416431 Meso_2876 polysaccharide deacetylase) Oant_0391 (416428..417672 Meso_2877 glycosyl transferase group 1) and Oant_0392 (417672..418799 Meso_2878 glycosyl transferase group 1) which in

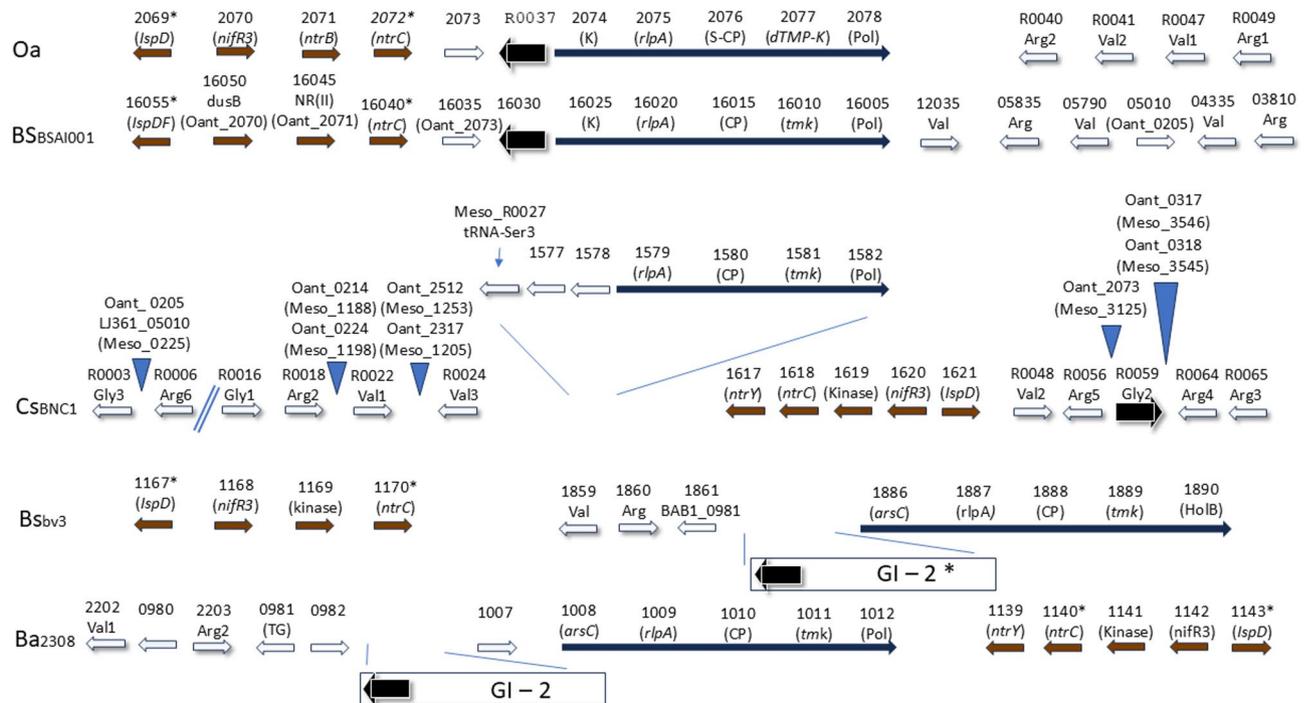


Fig. 3. The genetic maps of the $tRNA^{Gly}$ ($attB$) DNA region (The maps are not to scale, sequences are aligned). Top-*O. anthropi* ATCC 49188 (Locus tag = Oant_). Second-*Brucella* sp. JSBI001 (Locus tag = LJ361_). Third-*Chelativorans* sp. BNC1 (Locus tag = Meso_). Fourth-*B. suis* bv. 3, strain 686 (Locus tag = DK67_). Fifth-*B. abortus* 2308 (Locus tag = BAB1_). The dark blue long arrows stand for the conserved gene cluster which is shared by all strains (elaborated in Fig. 2); The black shadowed arrows stand for the $tRNA^{Gly}$ gene; The brown arrows stand for the nitrogen metabolizing genes, each indicated in parenthesis; The blank shadowed arrows stand for the tRNA encoding genes (note the specific tRNA – Ser which precedes upstream the conserved DNA unit in *Chelativorans* sp. BNC1). The blue triangle inserts between the genes indicate the location where the *O. anthropi* tagged genes merge with the *Chelativorans* sp. BNC1 genes. The arrows stand for the specific genes and the heads show the direction of the gene transcription. The numbers above the arrows depict the specific locus tags per a strain. The shadowed arrows stand for the specific tRNA CDS. The distances between the genes can be figured out from their locus tag numbers. To be noted: The *Mesorhizobium loti* conserved DNA unit and the nitrogen metabolizing gene cluster I downstream are preceded by a $tRNA^{Ser}$ at the 5'-side (Fig. 4). The *Chelativorans* sp. BNC1 DNA, between Meso_R0027 at the 5'-side (vertical arrow) and Meso_1621 at the 3'-side, has a similar gene synteny as the one in *M. loti*. A similar gene synteny is also shown in *B. suis* bv. 3, strain 686 and *B. abortus* strain 2308, however, the “arsenate reductase and related ($arsC$)” and GI-2 substitutes for the $tRNA^{Ser}$ at the 5'-side. The *O. anthropi* ATCC 49188 and *Brucella* sp. JSBI001 DNA units lack GI-2 and their conserved DNA units begin with “diacylglycerol kinase catalytic region (K)” rather than $arsC$, wherein the $tRNA^{Ser}$ is substituted for $tRNA^{Gly}$ and upstream a reversely oriented nitrogen metabolizing gene cluster. Note that the iconic $tRNA^{Gly}$ is positioned in *Chelativorans* sp. BNC1 in a reverse orientation further downstream. Exceptionals are: *B. inopinata* BO2 in which the scar is positioned in chromosome 2 but the nitrogen metabolizing gene cluster is located in chromosome 1; *Brucella* sp. 09RB8910 (ChrI), *Brucella* sp. 09RB8471 (ChrII) and *Brucella* sp. 458 (ChrI) depict similarity to the *B. suis* bv. 3, strain 686 chromosomal architecture (data not shown).

Chelativorans sp. BNC1 lay in vicinity to R0059 ($tRNA^{Gly}$) (Fig. 3) and further downstream Oant_0608 (638790..639557 Meso_1375 a glycosyl transferase family 25 protein) and Oant_0609 (639554..640813 mes:Meso_1374 an O-antigen polymerase) which link *Ochrobactrum* O-antigen and polysaccharide metabolizing enzymes to *Chelativorans* sp. BNC1.

Figure 3 compares the genetic map of the *Chelativorans* sp. BNC1 chromosome at the $tRNA^{Gly}$ locus to that of the *O. anthropi* ATCC 4188 and *Brucella* sp. JSBI001 chromosomes, on one hand, and the *B. suis* bv. 3, strain 686, and *B. abortus* strain 2308 chromosomes, on the other. It can be seen that the architecture of the *Brucella* and *Chelativorans* sp. BNC1 chromosomes differs from the *Ochrobactrum* strains, but the $tRNA^{Gly}$ which comprises the first gene in *Ochrobactrum*, and is part of GI-2 in *Brucella*, is substituted for $tRNA^{Ser}$ in the *Chelativorans* sp. BNC1 chromosome. As further can be seen, superimposition of the *O. anthropi* ATCC 49188 Meso-associated genes on the *Chelativorans* sp. BNC1 genetic map establishes a DNA unit between R0018 ($tRNA^{Arg2}$) upstream and R0064 ($tRNA^{Arg4}$) downstream which has similar borderlines with the *O. anthropi* ATCC 49188 Oant_0214 and Oant_0224 upstream and Oant_0317 and Oant_0318 downstream, and the nitrogen metabolizing genes reversely organized in the chromosome on the other side of the conserved gene region. Moreover, the same DNA

unit is depicted within a shorter *O. anthropi* ATCC 49188 DNA stretch, between Oant_2317 and Oant_2512 upstream and Oant_2073 downstream, in line with the organization of the *Brucella* sp. JSBI001 genes between LJ361_16050 (Oant_2070) upstream and LJ361_05010 (Oant_0205) downstream.

As shown in Fig. 1, Oant_0205 is part of an *O. anthropi* ATCC 49188 DNA region which in Phage Pr corresponds to a *Chelativorans* sp. BNC1 meso_0225 associated DNA unit⁴⁸. As can be seen in Fig. 3, this positions the gene in the *Chelativorans* sp. BNC1 genetic map between R0003 (tRNA-Gly3) and R0006 (tRNA-Arg6), remote from the common genomic region which this strain shares with *Ochrobactrum* and *Brucella*. Moreover, this gene has sequence similarity to LJ361_05010 (data not shown) and Fig. 3 shows its relative position in the genetic map of strain JSBI001 between LJ361_05790 (tRNA-Val) and LJ361_04335 (tRNA-Val).

An *Mesorhizobium loti* integrative, conjugative element (ICE)

Existence of a similar nitrogen metabolizing gene cluster and a *tRNA^{Gly}* adjunct conserved DNA unit in the *Chelativorans* sp. BNC1, *Ochrobactrum* and *Brucella* genomes infers that this DNA was acquired horizontally from a common source. Bacteria of the family *Rhizobiaceae* could be donors of such a gene cargo as they are nitrogen fixing endosymbionts of the *legume* roots which supplement the plants with reduced forms of nitrogen. In focus are *Mesorhizobium* spp. which harbor a mobile integrative, conjugative element (ICE) in their chromosome, and the element can form a conjugative integrative circular DNA which transfers horizontally to a broad spectrum of recipient cells, in which several forms are constructed as a composite multipartite ICE⁵².

An incomplete genome sequence of the *Mesorhizobium loti* strain R7ANSxAA22, collected in 62 annotated contigs, was recently published (<https://www.ncbi.nlm.nih.gov/Traces/wgs/NSFP01?display=contigs>). Each of the contigs was opened and searched individually for the presence of the nitrogen metabolizing gene cluster as well as the *Chelativorans* sp. BNC1, *Ochrobactrum* and *Brucella* DNAs *tRNA^{Gly}* adjunct conserved DNAs (Fig. 3). As can be seen in Fig. 4, a contiguous sequence which comprises the complete DNA unit was solely identified in Contig 15, e.g., Panel A includes the four genes that construct the conserved DNA unit and Panel B downstream includes the nitrogen metabolizing genes, organized in three sub-clusters, respectively. However, comparative analysis of the four genomes identified distinct difference in the gene content and their organization in the chromosomes. As can be seen, the *Ochrobactrum* and *Brucella* iconic *tRNA^{Gly}* is substituted for *tRNA^{Ser}* in the *M. loti* and the *Chelativorans* sp. BNC1 DNAs. Further, the *M. loti* nitrogen metabolizing gene cluster is much

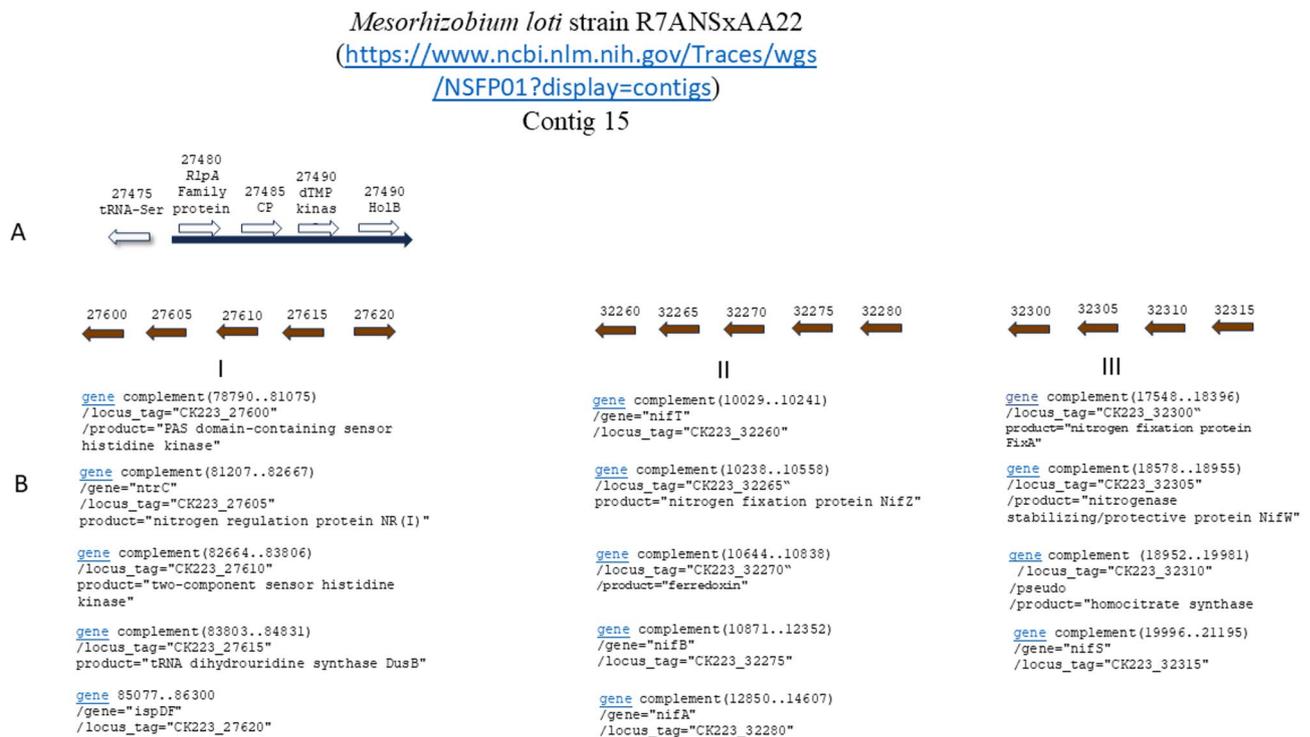


Fig. 4. The genetic map of the *Mesorhizobium loti* strain R7ANSxAA22 (Contig 15) ICE region (The map is not to scale). **(A)** A schematic representation of the conserved DNA unit (see Fig. 2). **(B)** A schematic representation of the nitrogen metabolizing gene cluster, as well as the sequence position, locus tags and gene products (see Fig. 3). The nitrogen metabolizing gene cluster includes three sub-clusters, e.g., I, II, and III, and the CDS are shown in the columns below. The arrows stand for the specific genes and the heads show the direction of the gene transcription. The numbers above the arrows depict the specific locus tags. The brown arrows stand for the nitrogen metabolizing genes, each indicated in parenthesis. Note that a *tRNA^{Ser}* is positioned upstream the conserved DNA unit, which in *Ochrobactrum* is substituted for “diacylglycerol kinase catalytic region (K)” and *tRNA^{Gly}* which upstream is bound to the reversely oriented nitrogen metabolizing gene cluster and in *Brucella* is substituted for arsenate reductase and related” (*arsC*) and GI-2 (e.g., *tRNA^{Gly}*).

larger and includes three sub-clusters rather than the Cluster 1 which comprises the genes in the chromosomes of the other three genera. Finally, the architecture of the *Ochrobactrum* chromosome varies in the location of the nitrogen metabolizing gene cluster upstream *tRNA^{Gly}* which, in the other three genera, is located in a reverse orientation further downstream.

Identification of the site where ChrI and ChrII divided

To pursue the question- the chromosome of strain JSBI001 was used as a reference genome to inquire about the gene synteny and the collective gene content of the bi-chromosomal organisms. In accordance- the DNA region in the chromosome of strain JSBI001- between LJ361_12035 and LJ361_05835- was predicted to include the entire gene content of chromosome 2. This hypothesis was tested by randomly selecting genes in this region- at around 100 ORFs apart- and BLASTn analyzing their sequences against *O. anthropi* ATCC 49-188 (taxid:439-375)- finding them all to be included in *O. anthropi* ATCC 49-188 ChrII (data not shown). Similarly- the same analysis was conducted on the *B. suis* bv. 3- strain 686 chromosome- between DK67_1170 and DK67_1859 or DK67_1167 and DK67_1886- against *Brucella abortus* 2308 (taxid:359-391)- finding all the genes located in the ChrI of the *B. abortus* strain 2308. Missing a hint where the genes that curate ChrII are located in the chromosome of *B. suis* bv. 3- strain 686- an opposite search was conducted- BLASTn comparing the *B. abortus* 2308 ChrII genes against *Brucella suis* bv. 3 str. 686 (taxid:520-487)- finding them all hit the *B. suis* bv. 3- strain 686 ChrII- Sequence ID: CP007718.1 (Indeed- GenBank assigns two chromosomes to this strain- e.g.- GB: CP007719 and GB: CP007718). This result inexplicitly shows that the *tRNA^{Gly}* site- which establishes *attB* in *Brucella* species- presents the circulation site where ChrI was closed and ChrII developed as a second replicon.

Conclusion remarks

The study of the brucellaphage Pr genome revealed sequence similarities to a specific DNA region in the *O. anthropi* ATCC 49188 chromosome 1 and the *Brucella* GI-2, respectively. This inferred an evolutionary path that GI -2 underwent in the *Brucella* genome showing evidence that this element integrated into a common DNA region in the chromosomes of the two clades, however, at a reverse position and therefore establishing distinct chromosomal architectures at this region. Adding *Mesorhizobium loti* to the study helped identify that the common DNA unit which characterizes the *Ochrobactrum* and *Brucella* DNAs originated from a mobile integrative, conjugative element (ICE). It is concluded that following the event of the ICE integration into the chromosome of an archetype strain *Brucella* and *Ochrobactrum* diverged, and GI -2 was then integrated into the ICE of the evolving *Brucella* lineage.

Discussion

Ochrobactrum is polyphyletic in comparison to *Brucella*, which even after the inclusion of the atypical strains still forms a monophyletic taxon⁵³. Within *Ochrobactrum* group A, the analysis shows that *O. anthropi* and *O. intermedium* among five other species establish sister taxa to *Brucella*, in support of the suggestion that the genus *Ochrobactrum* can be unified with the genus *Brucella*³⁰. This, however, was rebutted, in line with the argument that zoonosis and pathogenicity to humans is critical to the genus definition³¹, and as a matter of fact, inflicts on the national regulatory policies and reporting in some countries³². Moreover, the natural affiliation of core *Brucella* species to species specific animal hosts, in which they cause abortion in the last trimester of pregnancy, as well as their sensitivity to brucellaphages in a species-specific pattern, establishes an even more accurate criterion how *Brucella* are determined.

A study of the brucellaphage Iz1 genome identified short sequence homologies between the phage and the *O. anthropi* ATCC 491188 and the *Brucella* GI-2 DNAs⁵⁴. Months later, the publication of the genome sequence of Phage Pr, which is a duplicate of Phage Iz1, confirmed the sequencing⁴⁸. This afforded an in-deep analysis of the sequences which determine the the overlap between the phage and the GI-2 element at two specific phage related DNA regions, the *Chelativorans* sp. BNC1 DNA region, and a defined novel DNA region which is derived from the chromosome 1 of the *Ochrobactrum* strain, respectively (Fig. 1).

Analysis of scar, which determines the borderlines of the GI-2 element in the *Brucella* genome⁴⁹, identified several molecular patterns which enabled the integration and processing of the GI-2 element in the *Brucella* genome. The first depicts sequence editing which led the division of the *B. melitensis* 16 M BMEI0993 CDS into two distinguished *B. abortus* 2308 CDS, BAB1_1006 which bears the phage *attP* sequence and BAB1_1007, which is positioned out of the GI-2 element, respectively. This implies that at origin the phage *attP* sequence was larger than first predicted⁴⁹ (Fig. 2). The second shows how GI-2 excised the evolving genome via an integration/excision phage integrase precise activity which led the development of a consensus scar *attP* sequence in the chromosome of the stable *B. ovis* rough form⁴⁹. Thirdly, the modified sequence of the *B. inopinata* BO1 scar possibly indicates that illegitimate recombination led the processing of the GI-2 integration into the genome of this *Brucella* (Fig. 2, Table S1). Forth, GI-2 in *B. suis* bv. 3, strain 686, includes novel genes which in *B. abortus* strain 2308 are missing but their sequences are intergenic or partially overlapping with the *B. abortus* CDSs, and thus indicate that GI-2 in origin was larger and a more complexed element than the contemporary known structures, e.g., see for *B. inopinata* BO3 below as an example (Table S1). Fifth-*B. inopinata* BO3 lacks the iconic phage integrase at the 5' site and is missing *attP* at the 3'-end, whereas the genes downstream *tRNA^{Gly}* (*attB*) are novel, suggesting that GI-2 at origin was larger than the contemporary GI-2 elements and speciation as well as gene reduction led their individual evolution later (Fig. 2, Table S1). Sixth-truncated GI-2 structures in *Brucella* sp. 09RB8471, *Brucella* sp. 10RB9215 isolate BR10RB9215, and *Brucella* sp. NM4 (Table S1) further support the notion that the GI-2 at origin was different. Finally, the presence of IS sequences within GI-2 elements, e.g., *IS711* (most *Brucella*), *IS481* (*B. inopinata* BO3), *IS4* (*Brucella* sp. 2002734562) and *IS5* (*Brucella* sp. 2594,

Brucella sp. 1315) establishes an additional mode of gene acquisition and sequence modification, which could involve gene inactivation and genome enlargement, respectively.

Jumas-Bilak et al. hypothesized that homologous recombination at the ribosomal RNA (*rrna*) gene cluster site led chromosomal schism and development of the bi-chromosomal *Brucella*⁵⁰. However, chromosomal schism was refuted because the distribution frequency of the core genes in comparison to the metabolizing genes in the second chromosome cannot be accounted to a random schism mechanism and better fits the translocation of specific DNA units from the larger chromosome to a growing megaplasmid replicon⁵⁵. Here, bringing evidence that *Brucella* sp. JSBI001 is misplaced in the *Brucella* tree and rather conforms to *Ochrobactrum* (Table S1), broadens the spectrum in which at least one strain in the genus has a monopartite genome. In this regard, accepting the theory that the two chromosomes of the contemporary strains were fused in a saved genomic pattern, in two organisms, is too farfetched. Thence, the counter possibility that the bi-chromosomal *Brucella* evolved following chromosomal schism is strengthened, and data provided in this paper further support this notion showing that the *rrna* DNA locus resides at around 1000 CDS distance from the *tRNA^{Gly}* site, e.g., between DK67_687 (729304..730776) and DK67_693 (complement (735257..736648)).

Whole genome sequencing (WGS) of the *B. suis* genome indicated that it has an extensive gene synteny with the *Mesorhizobium loti* plant symbiont chromosome⁵⁶. *M. loti* strains are shown to include a circa 500 kb integrative, conjugative element (ICE), which encodes nitrogen metabolizing genes and regulatory functions⁵⁷. Therefore, existence of the nitrogen metabolizing gene cluster in a conserved gene synteny in the three chromosomes of *Chelativorans* sp. BNC1, *Ochrobactrum* and *Brucella*, could suggest that this DNA unit circulated among the *alphaproteobacteria* in the early history of their development. Indeed, this study shows that a similar element is found in the chromosome of the *M. loti* strain R7ANSxAA22, however, having the nitrogen metabolizing gene cluster being composed of three rather than the sole Cluster 1, and the *tRNA^{Ser}* which was shown to be common between the *M. loti* and the *Chelativorans* sp.BNC1 chromosomes being substituted for *tRNA^{Gly}* in the chromosomes of the *Ochrobactrum* and the *Brucella* (Fig. 4). Existence of a tri-partite ICE among symbiotic mesorhizobia throws light how such differences could evolve via recombinational driven chromosomal rearrangements later⁵⁸.

Having that said, there is a second subject which needs to be addressed in this chapter. Comparative examination of the architectures of the ICEs of the four organisms reveals that the *Ochrobactrum* ICE stands alone in relationships to the other three, regarding the location, at a reversely oriented gene pattern, of the nitrogen metabolizing gene cluster upstream *tRNA^{Gly}*. This divergence can best be explained if the *M. loti* ICE dominated during the ancient HGT events and the *Ochrobactrum* ICE evolved following a second translocation of its components later. Such a diversion infers that the *Ochrobactrum* archetype strain evolved as a daughter cell of the evolving *Brucella* lineage.

Phage Pr sequences teach us that GI-2 (and very likely Phage Pr) evolved progressively following recombinational events in a precursor host cell from which contemporary *Brucella* emerged (Fig. 1). Existence of sequence variations at the *Brucella* GI-2 DNA region between BAB1_0995 and BAB1_0997 (Table S1) further establishes that *IS711* was a hot spot where these recombinational events occurred. It is hypothesized that the 5'-end of this element constituted a DNA region in the *Chelativorans* sp. BNC1 chromosome (Fig. 3), and the 3'-end which bears *wboA* and *wboB* (Fig. 1) was integrated later (Fig. 4). Similarly, yet in the state of the *B. suis* bv. 3, strain 686 monopartite genome bearing organism, the *tRNA^{Ser}* was substituted for *tRNA^{Gly}* simultaneously with the translocation of the tRNA-Arg and tRNA-Val DNA region (Fig. 4). It can be envisioned, that additional DNA rearrangements at the site led chromosomal schism from which bi-chromosomal *Ochrobactrum* and *Brucella* species evolved. Finally, the development of the *B. melitensis*/*B. abortus* clade can be envisioned via explosive radiation and speciation into professional intracellular organisms, in adaptation to specific eco - systems⁵⁹.

Presence of *attP* in BAB1_1006, which in *B. melitensis* BME10993 establishes a composite gene (Fig. 2), echoes a *Vibrio cholera* XST *attP* sequence which composes an intergenic sequence in the 5' end of *prfC*⁶⁰. The fact that Phage Pr has a *Listonella* DNA in its genome could imply that a Phage Pr-like replicon imported a *V. cholera* XST-like element, which included the 3'-half part of GI-2, to a LJ361_05010-like gene (Fig. 3). In a similar event, de novo synthesis of the Phage Pr *O. anthropi* ATCC 49188 related DNA fragment (Fig. 1, Table 1) complemented the development of a functional lytic Phage Pr. Supplemented with sLPS and an OMP phage receptor⁶¹, emerging *Brucella* cells then provided the necessary niche for the propagation of a lytic Phage Pr progenitor (Fig. 5).

The species specific natural affiliation of core *Brucella* spp. to a natural host is a hallmark of this genus, and humans are only secondary passive hosts to these bacteria. Therefore, controlling animal brucellosis via vaccination⁶², housing animals in confinement⁶³ and controlling their movement play a major role in reducing human cases. Whereas *B. melitensis* exceeds with special virulence²⁹, *B. abortus* and *B. suis* pose similar risks, and *B. canis*, a pet's transmitted organism⁶⁴ and marine mammals' strains which closely relate to core *Brucella* correlate with the less chances of contracting the disease^{65,66}. In contrast, *B. neotomae*^{67,68}, and *B. ovis*^{7,69} are rarely associated with humans, overruling their zoonotic relevance.

Unclassified *Brucella* such as *Brucella* sp. 2594, *Brucella* sp. 1315 as well as *Brucella* sp. 10RB9215 isolate BR10RB9215 are frog isolates²⁰ and not yet associated with human cases. Does the fact that GI-2 is modified (Table S1) or their being pathogens of animal hosts which humans do not associate explain lack of human cases? *Brucella* sp. 6810 is a human isolate but its genome lacks GI-2 and *Brucella* sp. 09RB8471 (African bullfrog⁷⁰), similar to *B. inopinata* BO3 which is the first published amphibian strain associated with a human case²¹, has novel GI-2 and therefore cannot accomplish the complete path of OPS polymerization (Table S1). Accordingly, despite showing evidence regarding the role GI-2 plays in *Brucella* virulence, the understanding is that a criterium for presence or lack of GI-2 does not suffice to determine their risks to humans, in comparison to the *Ochrobactrum* pathogens which, nonetheless, their association to animal sources is scarce⁷¹.

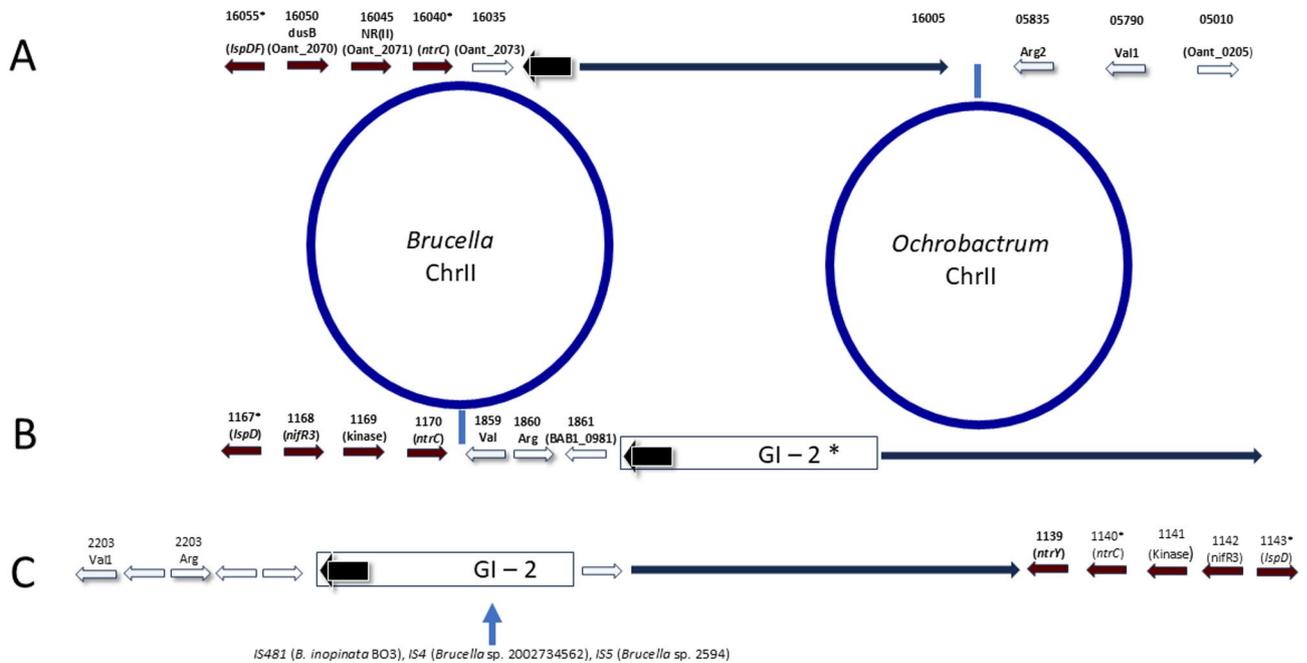


Fig. 5. A schematic model that elaborates how bi-chromosomal *Ochrobactrum* and *Brucella* evolved (The figure is not to scale). **(A)** A *Brucella* sp. JSBI001—like strain establishes the prototypic monopartite genome bearing precursor strain from which *Ochrobactrum* evolved. The blue circle represents ChrII in the bi-chromosomal strains. **(B)** The *B. suis* bv.3, strain 686, prototypic single chromosome of the evolving *Brucella*. The blue circle represents ChrII in the bi-chromosomal *Brucella*, which in this strain is annotated in the GenBank as GB: CP007718. The architecture of the chromosome and the gene synteny at this site is similar to the *M. loti* and the *Chelativorans* sp. BNC1 DNA loci, but the *tRNA^{Ser}* is substituted for *arsC* and GI-2. The specific overlapping of the nitrogen metabolizing gene cluster with the *Brucella* sp. JSBI001 DNA is deceptive because of the shortages in the chromosomal size following gene reduction. **(C)** The *B. abortus* strain 2308 chromosome 1 architecture and the corresponding gene synteny which establishes the prototypic form of the *M. loti* ICE site. The *tRNA^{Ser}* is substituted for *arsC* and GI-2. The insertion of IS481 (*B. inopinata* BO3), IS4 (*Brucella* sp. 2002734562) and IS5 (*Brucella* sp. 2594) led specific modifications in the GI-2 of these strains. The arrows stand for the specific genes and the heads show the direction of the gene transcription. The numbers above the arrows depict the specific locus tags. The shadowed arrows stand for the specific tRNA CDS. The distances between the genes can be figured out from their locus tag numbers. The black shadowed arrows stand for the tRNA-Gly2 gene; The brown arrows stand for the nitrogen metabolizing genes, each indicated in parenthesis; The blank shadowed arrows stand for the specific tRNA CDS. Notably, an inaccurate schism of the *Ochrobactrum* or the *Brucella* chromosome led the development of cognate strains in which the ICE is included in ChrII.

Materials and methods

List of organisms and sequence IDs

Brucella and *Ochrobactrum* strains and their source of isolation (copied from the published genomes), as well as the corresponding chromosomal sequence IDs, are shown in Table S1 and Table S2, respectively.

Comparative BLASTn sequence analyses

Sequence similarity between Phage Pr and PAIB1 was inferred by BLASTn analysis of the complete Phage Pr genome (JN939332.1) against *Brucella abortus* 2308 (taxid:359-391). A similar query was conducted against *Ochrobactrum anthropi* ATCC 49-188 (taxid:439-375) in search for sequence identity between Phage Pr and *Ochrobactrum*.

Identification of *attB* and *attP* sequences

The excision of PAIB1 from the chromosome of *Brucella melitensis* biovar Ovis 63/290 establishes a concrete scar sequence which was published by Vizcaíno et al. 2004 (Sequence ID: AY484543.1)⁴⁶. 264 nucleotides from this sequence include BAB1_1007 and BAB1_1006 (phage attachment sequence- *attP*) at the 5'-end- and *tRNA^{Gly}* (bacterial attachment sequence- *attB*) at the 3'-end- respectively. The sequence was submitted to BLASTn queries against several taxa IDs: *Brucella abortus* 2308 (taxid:359-391); *Brucella melitensis* bv. 1 str. 16 M (taxid:224-914); *Brucella suis* bv. 3 str. 686 (taxid:520-487); *Brucella inopinata* Schholz et al. 2010 (taxid:1-218-315); and unclassified *Brucella* (taxid:2-632-610). To find scar in the *Ochrobactrum* species- the scar sequence was similarly BLASTn queried against: *Ochrobactrum anthropi* ATCC 49-188 (taxid:439-375); *Brucella anthropi* ATCC 49-188 (taxid:439-375) and *Brucella intermedia* (taxid:94-625).

tRNA^{Gly} chromosomal location and GI-2 gene contents

BAB1_1007 is the first gene outside GI-2 and *attP* finds BAB1_1006 which establishes its 5' end boundaries. BAB1_2204 (*tRNA^{Gly}*) is the bacterial attachment site (*attB*) where GI-2 integrated to the chromosome. Together, these sequences establish the GI-2 boundaries in the *Brucella* chromosome. This enabled to identify each of the genes which are included in the GI-2 element and then BLASTn query their CDS against *Brucella abortus* 2308 (taxid:359391). Thus, hits in the *B. abortus* 2308 chromosome, Sequence ID: AM040264.1, identified sequence similarities to the specific PAIB1 genes. As *Ochrobactrum* species lack GI-2, scar queries identified the site where *tRNA^{Gly}* resides, and the linkage to the conserved DNA unit confirmed its identity. In addition, sequence hits found the specific chromosome in which *tRNA^{Gly}* or GI-2 were found, as detailed in Table S1 and Table S2.

Gene synteny between *Brucella* and *Ochrobactrum*

Genes that flank GI-2 in *Brucella* or *tRNA^{Gly}* in *Ochrobactrum* at the 3' end were found via chromosomal walking along the DNA. Each CDS was then BLASTn analyzed twice, once against *Brucella abortus* 2308 (taxid:359391) and secondly against *Ochrobactrum anthropi* ATCC 49188 (taxid:439375), thus confirming or refuting sequence similarities between the two organisms.

Identification of an integrative, conjugative element (ICE) in *Mesorhizobium loti*

The incomplete genome of *Mesorhizobium loti* strain R7ANSxAA22 (<https://www.ncbi.nlm.nih.gov/Traces/wgs/NSFP01?display=contigs>, NSFP0000000.1 *Mesorhizobium loti*) has been divided into 62 separate contigs. Each of the contigs was opened individually and searched for the presence of the nitrogen metabolizing genes and adjunct conserved genes in a cluster mode using their acronyms.

Data availability

Sequence IDs of the relevant chromosomes of all the investigated *Brucella* and *Ochrobactrum* strains are provided in Tables S1 and S2, respectively. *Brucella* genomes can be accessed in addition in the two links below: <https://www.ncbi.nlm.nih.gov/genome/browse#!/overview/Brucella> <https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=234>; Informative accession numbers are provided in Scholz et al. (2016). The *M. loti* strain R7ANSxAA22 genome is provided at: (<https://www.ncbi.nlm.nih.gov/Traces/wgs/NSFP01?display=contigs>).

Received: 27 February 2025; Accepted: 2 June 2025

Published online: 06 June 2025

References

- Bruce, D. A note on the discovery of a microorganism in Malta fever. *Practitioner* **39**, 161–170 (1887).
- Wyatt, H. V. How Themistocles Zammit found Malta Fever (brucellosis) to be transmitted by the milk of goats. *J. R. Soc. Med.* **98**, 451–454. <https://doi.org/10.1258/jrsm.98.10.451> (2005).
- Bang, B. Die aetiologie des seuchenhaften ('infectiosen') *Verwerfens*. *Z. Tiermed.* **1**, 1 (1897).
- Traum, I. Immature and hairless pigs. Report of the Department of Agriculture for the year ended June 30, 1914 (Washington) (1914).
- Evans, A. C. Further studies on *Bacterium abortus* and related bacteria. *J. Infect. Dis.* **22**, 580–593 (1918).
- Meyer, K. F. & Shaw, E. B. A comparison of the morphologic, cultural and biochemical characteristics of *B. abortus* and *B. melitensis*. Studies on the genus *Brucella* Nov. Gen. I. *J. Infect. Dis.* **27**, 173–184. <https://doi.org/10.1093/infdis/27.3.173> (1920).
- Huddleson, I. F. The differentiation of the species *Brucella*. *Michigan Agricultural Experimental Station Technical Bulletin* (No. 100), 1–16 (1929).
- Banai, M. & Corbel, M. Taxonomy of *Brucella*. *Open Vet. Sci. J.* **4**, 85–101. <https://doi.org/10.2174/1874318801004010085> (2010).
- Corbel, M. J. & Morgan, W. J. B. Proposal for minimal standards for descriptions of new species and biotypes of the genus *Brucella*. *Int. J. Syst. Evol. Microbiol.* **25**, 83–89. <https://doi.org/10.1099/00207713-25-2-243> (1975).
- Verger, J. M., Grimont, F., Grimont, P. A. D. & Grayon, M. *Brucella*, a monospecific genus as shown by deoxyribonucleic acid hybridization. *Int. J. Syst. Evol. Microbiol.* **35**, 292–295 (1985).
- Ficht, T. *Brucella* taxonomy and evolution. *Future Microbiol.* **5**, 859–866 (2010).
- Tiller, R. V. et al. Characterization of novel *Brucella* strains originating from wild native rodent species in North Queensland, Australia. *Appl. Environ. Microbiol.* **76**, 5837–5845. <https://doi.org/10.1128/AEM.00620-10.16> (2010).
- Brangsch, H., Horstkotte, M. A. & Melzer, F. Genotypic peculiarities of a human brucellosis case caused by *Brucella suis* biovar 5. *Sci. Rep.* **13**, 16586. <https://doi.org/10.1038/s41598-023-43570-4> (2023).
- Garin-Bastuji, B. et al. Examination of taxonomic uncertainties surrounding *Brucella abortus* bv. 7 by phenotypic and molecular approaches. *Appl. Environ. Microbiol.* **80**, 1570. <https://doi.org/10.1128/AEM.03755-13> (2014).
- Suárez-Esquivel, M. et al. *Brucella* genetic variability in wildlife marine mammals populations relates to host preference and ocean distribution. *Genome Biol. Evol.* **9**, 1901–1912. <https://doi.org/10.1093/gbe/evx137> (2017).
- Scholz, H. C. et al. *Brucella vulpis* sp. nov., isolated from mandibular lymph nodes of red foxes (*Vulpes vulpes*). *Int. J. Syst. Evol. Microbiol.* **66**, 2090–2098. <https://doi.org/10.1099/ijsem.0.000998> (2016).
- Whatmore, A. M. et al. *Brucella papionis* sp. Nov., isolated from baboons (*Papio* spp.). *Int. J. Syst. Evol. Microbiol.* **64**, 4120–4128. <https://doi.org/10.1099/ijss.0.065482-0> (2014).
- Scholz, H. C. et al. *Brucella microti* sp. Nov., isolated from the common vole *Microtus arvalis*. *Int. J. Syst. Evol. Microbiol.* **58**(2), 375–382. <https://doi.org/10.1099/ijss.0.65356-0> (2008).
- Whatmore, A. M. et al. Marine mammal *Brucella* genotype associated with zoonotic infection. *Emerg. Infect. Dis.* **14**, 517–518. <https://doi.org/10.3201/eid1403.070829> (2008).
- Scholz, H. C. et al. The change of a medically important genus: Worldwide occurrence of genetically diverse novel *Brucella* species in exotic frogs. *PLoS ONE* **11**, e0168872. <https://doi.org/10.1371/journal.pone.0168872> (2016).
- Rouziq, N. et al. First case of brucellosis caused by an amphibian-like *Brucella*. *Clin. Infect. Dis.* **72**, e404–e407 (2021).
- Soler-Lloréns, P. F. et al. A *Brucella* spp. isolate from a Pac-Man frog (*Ceratophrys ornata*) reveals characteristics departing from classical *Brucellae*. *Front. Cell. Infect. Microbiol.* **6**, 116. <https://doi.org/10.3389/fcimb.2016.00116> (2016).
- Leclercq, S. O., Cloeckert, A. & Zygmunt, M. S. Taxonomic organization of the family *Brucellaceae* based on a phylogenomic approach. *Front. Microbiol.* **10**, 3083. <https://doi.org/10.3389/fmicb.2019.03083> (2020).
- Scholz, H. C. et al. *Brucella inopinata* sp. Nov., isolated from a breast implant infection. *Int. J. Syst. Evol. Microbiol.* **60**, 801–808. <https://doi.org/10.1099/ijss.0.011148-0> (2010).

25. Hammerl, J. A. et al. Analysis of the first temperate broad host range brucellaphage (BiPBO1) isolated from *B. inopinata*. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2016.00024> (2016).
26. Abdel-Glil, M. Y. et al. Core genome multilocus sequence typing scheme for improved characterization and epidemiological surveillance of pathogenic *Brucella*. *J. Clin. Microbiol.* **60**, e00311–e00322. <https://doi.org/10.1128/jcm.00311-22> (2022).
27. Bohlin, J. et al. Genomic comparisons of *Brucella* spp. and closely related bacteria using base compositional and proteome based methods. *BMC Evol. Biol.* **10**, 249 (2010).
28. Whatmore, A. M. & Foster, J. T. Emerging diversity and ongoing expansion of the genus *Brucella*. *Infect. Genet. Evol.* **92**, 104865. <https://doi.org/10.1016/j.meegid.2021.104865> (2021).
29. Al Dahouk, S. et al. Evaluation of *Brucella* MLVA typing for human brucellosis. *J. Microbiol. Methods* **69**, 137–145. <https://doi.org/10.1016/j.mimet.2006.12.015> (2007).
30. Hördt, A. et al. Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of *Alphaproteobacteria*. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2020.00468> (2020).
31. Moreno, E. et al. If you're not confused, you're not paying attention: *Ochrobactrum* is not *Brucella*. *J. Clin. Microbiol.* **61**, e0043823. <https://doi.org/10.1128/jcm.00438-23> (2023).
32. She, R. et al. On behalf of the American Society for Microbiology Clinical and Public Health Committee, Laboratory Practices Subcommittee. *Brucella* and *Ochrobactrum* Taxonomic Updates for Laboratories. <https://asm.org/Guideline/Brucella-and-Ochrobactrum-Taxonomic-Updates-for-La> (2023).
33. Moreno, E. et al. *Brucella abortus* 16S rRNA and lipid A reveal phylogenetic relationship with members of the alpha-2 subdivision of the class *Proteobacteria*. *J. Bacteriol.* **172**, 3569–3576 (1990).
34. Velasco, J. et al. Structural studies on the lipopolysaccharide from a rough strain of *Ochrobactrum anthropi* containing a 2,3-diamino-2,3-dideoxy-D-glucose disaccharide lipid A backbone. *Carbohydr. Res.* **306**, 283–290. [https://doi.org/10.1016/s0008-6215\(97\)10029-5](https://doi.org/10.1016/s0008-6215(97)10029-5) (1998).
35. Velasco, J. et al. *Brucella abortus* and its closest phylogenetic relative, *Ochrobactrum* spp., differ in outer membrane permeability and cationic peptide resistance. *Infect. Immun.* **68**, 3210–3218. <https://doi.org/10.1128/iai.68.6.3210-3218.2000> (2000).
36. Kubler-Kielb, J. & Vinogradov, E. The study of the core part and non-repeating elements of the O-antigen of *Brucella* lipopolysaccharide. *Carbohydr. Res.* **366**, 33–37. <https://doi.org/10.1016/j.carres.2012.11.004> (2013).
37. Cardoso, P. G., Macedo, G. C., Azevedo, V. & Oliveira, S. C. *Brucella* spp noncanonical LPS: Structure, biosynthesis, and interaction with host immune system. *Microb. Cell. Fact.* **5**, 13. <https://doi.org/10.1186/1475-2859-5-13> (2006).
38. Roop, R. M., Barton, I. S., Hoppersberger, D. & Martin, D. W. Uncovering the hidden credentials of *Brucella* virulence. *Microbiol. Mol. Biol. Rev.* <https://doi.org/10.1128/mmr.00021-19.10.1128/mmr.00021-19> (2021).
39. Mathur, S., Banai, M. & Cohen, D. Natural *Brucella melitensis* infection and Rev. 1 vaccination induce specific *Brucella* O-Polysaccharide antibodies involved in complement mediated *Brucella* cell killing. *Vaccines* **10**, 317. <https://doi.org/10.3390/vaccines10020317> (2022).
40. Kubler-Kielb, J. & Vinogradov, E. Reinvestigation of the structure of *Brucella* O-antigens. *Carbohydr. Res.* **378**, 144–147. <https://doi.org/10.1016/j.carres.2013.03.021> (2013).
41. Fontana, C. et al. Structural studies of lipopolysaccharide-defective mutants from *Brucella melitensis* identify a core oligosaccharide critical in virulence. *J. Biol. Chem.* **291**, 7727–7741. <https://doi.org/10.1074/jbc.M115.701540> (2016).
42. Godfroid, F. et al. Genetic organisation of the lipopolysaccharide O-antigen biosynthesis region of *Brucella melitensis* 16M (*wbk*). *Res. Microbiol.* **151**, 655–668. [https://doi.org/10.1016/S0923-2508\(00\)90130-X](https://doi.org/10.1016/S0923-2508(00)90130-X) (2000).
43. González, D. et al. Brucellosis vaccines: assessment of *Brucella melitensis* lipopolysaccharide rough mutants defective in core and O-polysaccharide synthesis and export. *PLoS ONE* **3**, e2760. <https://doi.org/10.1371/journal.pone.0002760> (2008).
44. Soler-Lloréns, P. et al. Mutants in the lipopolysaccharide of *Brucella ovis* are attenuated and protect against *B. ovis* infection in mice. *Vet. Res.* **45**, 72. <https://doi.org/10.1186/s13567-014-0072-0> (2014).
45. Chain, P. S. et al. Whole-genome analyses of speciation events in pathogenic *Brucellae*. *Infect. Immun.* **73**, 8353–8361. <https://doi.org/10.1128/IAI.73.12.8353-8361.2005> (2005).
46. Vizcaino, N., Caro-Hernández, P., Cloeckert, A. & Fernández-Lago, L. DNA polymorphism in the omp25/omp31 family of *Brucella* spp.: Identification of a 17-kb inversion in *Brucella cetaceae* and of a 151-kb genomic island, absent from *Brucella ovis*, related to the synthesis of smooth lipopolysaccharide. *Microbes Infect.* **6**, 821–834. <https://doi.org/10.1016/j.micinf.2004.04.009> (2004).
47. Wattam, A. R. et al. Comparative phylogenomics and evolution of the Brucellae reveal a path to virulence. *J. Bacteriol.* **196**, 920–930. <https://doi.org/10.1128/JB.01091-13> (2014).
48. Flores, V., López-Merino, A., Mendoza-Hernandez, G. & Guarneros, G. Comparative genomic analysis of two brucellaphages of distant origins. *Genomics* **99**, 233–240. <https://doi.org/10.1016/j.ygeno.2012.01.001> (2012).
49. Mancilla, M., López-Goñi, I., Moriyón, I. & Zárrega, A. M. Genomic island 2 is an unstable genetic element contributing to *Brucella* lipopolysaccharide spontaneous smooth-to-rough dissociation. *J. Bacteriol.* **192**, 6346–6351. <https://doi.org/10.1128/JB.00838-10> (2010).
50. Jumas-Bilak, E., Michaux-Charachon, S., Bourg, G., O'Callaghan, D. & Ramuz, M. Differences in chromosome number and genome rearrangements in the genus *Brucella*. *Mol. Microbiol.* **27**, 99–106. <https://doi.org/10.1046/j.1365-2958.1998.00661.x> (1988).
51. O'Callaghan, D. & Whatmore, A. M. *Brucella* genomics as we enter the multi-genome era. *Brief. Funct. Genomics* **10**, 334–341. <https://doi.org/10.1093/bfpg/elr026> (2011).
52. Colombi, E. et al. Comparative analysis of integrative and conjugative mobile genetic elements in the genus *Mesorhizobium*. *Microbial Genomics* **7**, 000657. <https://doi.org/10.1099/mgen.0.000657> (2021).
53. Ashford, R. T., Muchowski, J., Koylass, M., Scholz, H. C. & Whatmore, A. M. Application of whole genome sequencing and pan-family multi-locus sequence analysis to characterize relationships within the family *Brucellaceae*. *Front. Microbiol.* **11**, 1329. <https://doi.org/10.3389/fmicb.2020.01329> (2020).
54. Banai, M., Strada, V., Bardenstein, S., Ben-Asouli, I. & Osman, F. *Brucella* phage polynucleotides and uses thereof. US 8,722,411 B2. Granted 13.05.2014 (2012).
55. diCenzo, G. C. & Finan, T. M. The divided bacterial genome: Structure, function, and evolution. *Microbiol. Mol. Biol. Rev.* **81**, e00019. <https://doi.org/10.1128/MMBR.00019-17> (2017).
56. Paulsen, I. T. et al. The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *PNAS* **99**(20), 13148–13153. <https://doi.org/10.1073/pnas.192319099> (2002).
57. Sullivan, J. T. et al. Comparative sequence analysis of the Symbiosis Island of *Mesorhizobium loti* strain R7A. *J. Bacteriol.* **184**, 3086. <https://doi.org/10.1128/jb.184.11.3086-3095.2002> (2002).
58. Haskett, T. L. et al. Assembly and transfer of tripartite integrative and conjugative genetic elements. *PNAS* **113**(43), 12268–12273. <https://doi.org/10.1073/pnas.1613358113> (2016).
59. Suárez-Esquivel, M., Chaves-Olarte, E., Moreno, E. & Guzmán-Verri, C. *Brucella* genomics: Macro and micro evolution. *Int. J. Mol. Sci.* **21**, 7749. <https://doi.org/10.3390/ijms21207749> (2020).
60. Bioteau, A., Durand, R. & Burrus, V. Redefinition and unification of the SXT/R391 family of integrative and conjugative elements. *Appl. Environ. Microbiol.* **84**, e00485–e518. <https://doi.org/10.1128/AEM.00485-18> (2018).
61. Corbel, M. J. Isolation and partial characterization of a phage receptor from “*Brucella* 273 neotomae” 5K33. *Ann. Clav. Sci.* **19**, 131–142 (1977).

62. Elbehiry, A. et al. The development of diagnostic and vaccine strategies for early detection and control of human brucellosis, particularly in endemic areas. *Vaccines* **11**, 654. <https://doi.org/10.3390/vaccines11030654> (2023).
63. Olsen, S. C. & Tatum, F. M. Swine brucellosis: Current perspectives. *Vet. Med.* **8**, 1–12. <https://doi.org/10.2147/VMRR.S91360> (2017).
64. Wallach, J. C., Giambartolomei, G. H., Baldi, P. C. & Fossati, C. A. Human infection with M- strain of *Brucella canis*. *Emerg. Infect. Dis.* **10**, 146–148. <https://doi.org/10.3201/eid1001.020622> (2004).
65. Nymo, I. H., Tryland, M. & Godfroid, J. A review of *Brucella* infection in marine mammals, with special emphasis on *Brucella pinnipedialis* in the hooded seal (*Cystophora cristata*). *Vet. Res.* **42**, 93. <https://doi.org/10.1186/1297-9716-42-93> (2011).
66. Sohn, A. H. et al. Human neurobrucellosis with intracerebral granuloma caused by a marine mammal *Brucella* spp. *Emerg. Infect. Dis.* **9**, 485–488. <https://doi.org/10.3201/eid0904.020576> (2003).
67. Suárez-Esquivel, M. et al. *Brucella neotomae* infection in humans, Costa Rica. *Emerg. Infect. Dis.* **23**, 997–1000. <https://doi.org/10.3201/eid2306.162018> (2017). Erratum in: *Emerg. Infect. Dis.* **23**(8), 1435. <https://doi.org/10.3201/eid2308.C22308> (2017).
68. Kang, Y.-S., Brown, D. A. & Kirby, J. E. *Brucella neotomae* recapitulates attributes of zoonotic human disease in a murine infection model. *Infect. Immun.* **87**, e00255–e318. <https://doi.org/10.1128/IAI.00255-18> (2019).
69. Tsolis, R. M. et al. Genome degradation in *Brucella ovis* corresponds with narrowing of its host range and tissue tropism. *PLoS ONE* **4**, e5519. <https://doi.org/10.1371/journal.pone.0005519> (2009).
70. Al Dahouk, S. et al. *Brucella* spp. of amphibians comprise genomically diverse motile strains competent for replication in macrophages and survival in mammalian hosts. *Sci. Rep.* **7**, 44420. <https://doi.org/10.1038/srep44420> (2017).
71. Ryan, M. P. & Pembroke, J. T. The genus *Ochrobactrum* as major opportunistic pathogens. *Microorganisms* **8**, 1797. <https://doi.org/10.3390/microorganisms8111797> (2020).

Acknowledgements

Dr. Michael Corbel helped drafting an early version and then after, provided insightful comments throughout the writing. Research was partly funded by the Technical Support Working Group project #T-582 (2005–2009).

Author contributions

M.B. conceived the working hypothesis, carried out all the BLASTn analyses, conducted the manual curations of the chromosomes, drafted the manuscript and designed the figures and tables.

Declarations

Competing interests

Dr. Banai owns a patent, *Brucella* phage polynucleotides and uses thereof. US 8,722,411 B2. (Granted 13.05.2014). This patent addresses gene silencing in *Brucella*.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-05244-1>.

Correspondence and requests for materials should be addressed to M.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025