A carbonic anhydrase pseudogene sensitizes select *Brucella* lineages to low CO_2 partial pressure

Lydia M. Varesio[a], Jonathan W. Willett[b], Aretha Fiebig[b,c], Sean Crosson[b,c,#]

[a]Committee on Microbiology, University of Chicago, Chicago, Illinois, USA
[b]Department of Biochemistry and Molecular Biology, University of Chicago, Illinois, USA
[c]Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan, USA.

[#] To Whom Correspondence Should be Addressed: Michigan State University, Department of Microbiology and Molecular Genetics, 567 Wilson Rd., East Lansing, Michigan 48824. crosson4@msu.edu

Abstract

Brucella spp. are intracellular pathogens that cause a disease known as brucellosis. Though the genus is highly monomorphic at the genetic level, species have animal host preferences and some defining physiologic characteristics. Of note is the requirement for increased CO_2 tension to cultivate particular species, which confounded early efforts to isolate *B. abortus* from diseased cattle. Differences in the capacity to assimilate CO_2 are determined by mutations in the carbonic anhydrase gene, *bcaA*. Ancestral single nucleotide insertions in *bcaA* have resulted in frameshifted pseudogenes in *B. abortus* and *B. ovis* lineages, which underlie their inability to grow under the low CO_2 tension of a standard atmosphere. Incubation of wild-type *B. ovis* in air selects for mutations that "rescue" a functional *bcaA* reading frame, which enables growth under low CO_2 and enhances growth rate in high CO_2. Consistent with this result, we show that heterologous expression of functional *E. coli* carbonic anhydrases enables *B. ovis* growth in air. Growth of wild-type *B. ovis* is acutely sensitive to a reduction in CO_2 tension, while frame-rescued *B. ovis* mutants are completely insensitive to CO_2 shifts. Wild-type *B. ovis* initiates a gene expression program upon CO_2 downshift that resembles the stringent response and results in activation of its *virB* type IV secretion system. We conclude that loss-of-function single-nucleotide insertion mutations in *bcaA* sensitize the CO_2 response of *B. ovis* and *B. abortus* relative to other *Brucella* lineages. CO_2-dependent starvation and virulence gene expression programs in these species may influence transmission or infection in natural hosts.

Importance. *Brucella* spp. are highly related, but exhibit differences in animal host preference that must be determined by genome sequence differences. *B. ovis* and the majority of *B. abortus* strains require increased CO_2 tension to be cultivated *in vitro*, and harbor conserved insertional mutations in the carbonic anhydrase, *bcaA*, which underlie this trait. Mutants that grow in a standard atmosphere, first reported nearly a century ago, are easily selected in the laboratory. These mutants harbor varied indel polymorphisms in *bcaA* that restore its consensus reading frame and rescue its function. Loss of *bcaA* function has evolved independently in the *B. ovis* and *B. abortus* lineages, and results in strains with a dramatically increased sensitivity to CO_2 limitation.

Introduction

Bacteria of the genus *Brucella* are the etiologic agents of brucellosis, which is among the most common zoonotic diseases worldwide (1). *Brucella* spp. can infect a range of wild and livestock animals (2), but have a relatively narrow host range and varying zoonotic potential (3, 4) and pathogenicity in humans. *B. melitensis*, for example, is primarily a pathogen of sheep and goats and is often considered the most virulent species in humans (5). In contrast, human infections by the cattle pathogen, *B. abortus*, and the swine pathogen, *B. suis* (biovars 1, 3, and 4), are less frequent and often less pathogenic, though clinical differences between these species are difficult to discern (6).

Considering that genetic identity across the genus is 94-98% at the coding level (4, 7) differences in *Brucella* spp. host range, virulence, and zoonotic potential are notable. Comparative studies of genome content and the genetic requirements for growth in diverse environments may inform our understanding of differences in the physiology and infection biology of *Brucella* spp. A historically-significant physiologic difference between species is the requirement for CO_2 supplementation for axenic growth. In this manuscript, we provide a historical overview of CO_2 as a required

substrate for *Brucella* cultivation, describe a series of forward and reverse genetic and transcriptomic studies focused on understanding CO_2 metabolism in the ovine pathogen, *Brucella ovis*, and present a genus-level evolutionary analysis of a β-carbonic anhydrase gene (*bcaA*) that provides understanding of which *Brucella* strains require CO_2 supplementation for axenic growth.

Brucella *cultivation and CO_2: a brief history*

Bang and Stribolt pioneered the study of *B. abortus* growth in axenic culture, which was important in establishing its identity as the etiologic agent of contagious abortion in cattle (8). Of particular note in their original study was the growth pattern of *B. abortus* in solid shake tubes, in which they observed bacteria in two bands just below the surface of nutrient agar. Based on a series of experiments using different atmospheres of CO_2 and O_2, Bang concluded that oxygen and not carbon dioxide was the gaseous factor that governed growth, though he noted significant variability in growth zones in shake tubes (8). In the years following the discovery of Bang and Stribolt, reliable axenic cultivation of *B. abortus* from clinical specimens remained a major challenge (9).

Huddleson conducted a series of growth experiments in 1921, demonstrating that increasing the partial pressure of CO_2 (pCO_2) stimulated growth of *B. abortus* clinical isolates (10). A follow-up investigation provided additional evidence that CO_2 was required to support metabolism of *B. abortus*, and that the effect of CO_2 on growth was not a result of pH changes in the medium (11). G.S. Wilson later presented evidence that both oxygen and CO_2 *per se* are required for *B. abortus* growth, and that bicarbonate cannot substitute for CO_2 in the growth medium (12).

To assess the metabolic fate of CO_2, J.B. Wilson and colleagues grew multiple *B. abortus* strains in the presence of $^{14}CO_2$ and measured incorporation of the carbon into amino acids and nucleotides (13-15). Carbon-14 largely pooled in pyrimidines and in glycine: 90% of the carbon-14 detected in the amino acid pool was in glycine (15). Amino acids, purines, pyrimidines, various organic acids and cell hydrolysates were not able to substitute for CO_2 in their experiments (16). In the case of *B. abortus*, the incorporation of $^{14}CO_2$ into glycine and pyrimidines occurred at similar ratios in all strains tested, even those that did not require increased pCO_2 for growth (15). From these data, the authors concluded that fixation of CO_2 is a fundamental property of *B. abortus* strains.

Natural variation in the requirement for CO_2 supplementation

When investigating the cause of Malta fever, Bruce noted that *Micrococcus* (*Brucella*) *melitensis* appeared as colonies "on the surface" of nutrient agar in an air incubator (17). This stands in contrast with *Brucella abortus* isolates from cattle that did not grow on agar surfaces (8) without CO_2 supplementation (10). Indeed, variation in atmospheric requirements for growth of these bacteria clouded the fact that these early investigators were actually studying closely-related species (18). Pioneering work by Alice Evans united the etiologic agents of Malta fever in humans and contagious abortion in cattle into a single group (19) that she classified under the genus *Brucella* (20). By 1930, it was established that certain *Brucella* isolates required CO_2 for axenic cultivation, while others could be cultivated without increased pCO_2 (21). In the case of *B. abortus*, it was well known by the 1920's that the requirement for increased pCO_2 for growth was often spontaneously lost after a period of cultivation in the laboratory (11, 22). Buddle reported a similar phenomenon in an early description of *Brucella ovis* isolated from rams in New Zealand (23).

What, then, occurs when a *Brucella* isolate spontaneously loses the requirement for increased pCO_2 to grow? Early descriptions of "crops" of colonies that grew from clinical specimens on solid media after several days of incubation (11) are reminiscent of spontaneous genetic mutants that arise when bacteria are cultivated under selective conditions. Smith attempted to isolate genetic revertants of lab-adapted strains derived from these crops by exposing the cells to guinea pig tissue or re-infecting cows (11). However, he failed to identify strains that regained a requirement for CO_2 supplementation. Marr and Wilson later demonstrated that *B. abortus* spontaneously loses its requirement for CO_2 at a rate of 3×10^{-10} per generation (24), supporting a model whereby this phenotype is likely a result of mutation at a single site.

A model to investigate Brucella responses to CO_2 limitation

As a model to study CO_2 metabolism and gene regulation, we have selected *Brucella ovis*, a natural pathogen of sheep that has a well-documented requirement for increased pCO_2 to grow axenically (23, 25, 26). We report a forward genetic selection strategy to identify *B. ovis* mutants that grow spontaneously under standard atmospheric conditions, and have defined the genetic changes in these strains by whole genome sequencing. In agreement with recent results from Perez-Etayo and colleagues (27), we found that *B. ovis* mutants selected to grow in a

standard atmosphere carry a frameshift mutation at the 3' end of a β-carbonic anhydrase gene, *bcaA*$_{BOV}$. Carbonic anhydrases catalyze the reversible hydration of CO_2 to bicarbonate, an essential metabolic building block (28), and plasmid-based expression of the frameshifted *bcaA*$_{BOV}$ allele or of functional *Escherichia coli* carbonic anhydrases, *can* or *cynT*, are sufficient to enable growth of *B. ovis* in a standard atmosphere.

To better understand the selected *bcaA*$_{BOV}$ frameshift mutations, we conducted a comparative and evolutionary analysis of over 700 sequenced *Brucella* genomes, and identified polymorphisms in *B. ovis* and select *B. abortus* strains that result in non-functional, frameshifted *bcaA* pseudogenes. Our data provide evidence that *B. ovis* mutants selected to grow without CO_2 supplementation have "resurrected" the frameshifted *bcaA*$_{BOV}$ pseudogene to encode a functional BcaA protein sequence. The protein sequence of the resurrected BcaA$_{BOV}$ matches BcaA from *B. melitensis* and other *Brucella* species that do not require CO_2 supplementation. From these data, we conclude that the different atmospheric requirements for *Brucella* spp. cultivation noted by early investigators of *Brucella* physiology are a result of small insertion and deletion mutations in *bcaA*, and that functional revertants of *bcaA* pseudogenes are easily selected under CO_2 limitation. We further show that *1)* wild-type *B. ovis* is acutely sensitive to CO_2 limitation as a result of harboring a *bcaA* pseudogene, *2)* that CO_2 limitation triggers transcription of known starvation stress response and virulence genes including the *virB* type IV secretion system, and *3)* that gain-of-function frameshift mutations in *bcaA* render *B. ovis* insensitive to changes in CO_2 tension.

Materials and Methods

Bacterial strains and growth conditions

The bacterial strains used in this study are listed in **Data Set 4.** *Brucella ovis* was grown on Schaedler agar (Difco Laboratories) or on Tryptic Soy Agar (Difco Laboratories) plates supplemented with 5% defibrinated sheep blood (Quad Five) (SBA or TSBA respectively) at 37°C with 5% CO_2 when required. Liquid cultures were grown in Brucella Broth (BB) (Difco Laboratories) at 37°C with 5% CO_2 when required. 50 μg ml^{-1} of kanamycin (kan) was added when required. Expression strains were supplemented with 1 mM isopropy β-D-1-thiogalactopyranoside (IPTG) (GoldBio). All studies on *Brucella ovis* were performed following Biosafety Level 2 (BSL2) protocols.

Escherichia coli strains used for cloning were grown on lysogeny broth (LB) (Fisher Bioreagents) at 37°C and on LB plates with 1.5% agar (Fisher Bioreagents). Kan was added to a final concentration of 50 μg ml^{-1}. For conjugation, the diaminopimelic acid (DAP) auxotrophic *E. coli* strain WM3064 was grown with a final concentration of 300 μM DAP (Sigma-Aldrich).

Plasmid and Strain Construction

For the construction of allele replacement strains, external primers overlapping regions approximately 500 bp upstream or downstream flanking target genes of interest were PCR amplified with KOD Xtreme Hot Start Polymerase (Novagen) using either *Brucella ovis*, *Brucella ovis* mutant or *Brucella abortus* genomic DNA as a template. For deletion strains, internal primers were designed so that the 4-6 amino acids at the N-terminus and the C-terminus of the gene product would remain in frame to minimize polar effects. For strains with point or small mutations, the target gene was amplified using internal primers carrying the mutation of interest. The various fragments obtained were then joined by overlap extension PCR (OE-PCR) to generate null or mutant alleles using KOD. Insertion of fragments into the suicide plasmid pNPTS138 was attained either by restriction enzyme (RE) digestion and T4 ligase ligation or by Gibson assembly (New England Biolabs).

To build gene overexpression strains, genomic DNA from *Escherichia coli* MG1655, *Brucella ovis*, *Brucella abortus* or derivatives thereof was used as templates. Primers amplifying the gene of interest starting with the start codon and including the stop codon were amplified by PCR with KOD. Fragment insertion into the *lac*-inducible, low-copy pSRK plasmid was attained either by RE digestion and T4 ligation or by Gibson assembly. All insertion sequences into plasmids were confirmed by Sanger sequencing. Sequence-confirmed plasmids were transformed into *E. coli* Top10 (for plasmid maintenance), purified (ThermoFisher Scientific), and transformed into *E. coli* WM3064 (William Metcalf, University of Illinois) for conjugation into *Brucella ovis*. In the case of replicating plasmids, selection on SBA or TSBA kan plates (without DAP) allowed for isolation of single *B. ovis* clones containing the plasmid of interest. In the case of the pNPTS138 plasmid used to conduct gene replacement, cells from mating were first selected on SBA or TSBA kan plates. pNPTS138 also carries the *sacB* gene for counterselection. Strains harboring pNPTS138 integrants were outgrown in BB and spread on TSBA (or SBA) plates containing 5% (w/v) sucrose. This permitted selection of single clones in which a second recombination event that excised the plasmid had occurred. Clones were then screened by PCR with GoTaq® Green polymerase and the PCR products were Sanger sequenced (University of Chicago Comprehensive Cancer Center, DNA Sequencing & Genotyping Facility) to confirm the sequence of the gene deletion or replacement. See **Data Set 4** for primers, strains and plasmids.

DNA extraction, amplification and quantification

Genomic DNA was extracted following a standard guanidinium thiocyanate protocol. Briefly, strains were struck on SBA (or TSBA) agar plates and colonies were picked and cultivated in BB overnight. 1 ml of culture was spun at 12000 rpm for 20 s and the pellet washed with 0.5 ml of Phosphate-Buffered Saline (PBS). Pellet was

resuspended in 0.1 ml of TE buffer pH 8.0 (10mM Tris-HCl pH 8.0 (Fisher Bioreagents); 1 mM EDTA (Fisher Bioreagents)). TE buffer included ribonuclease A (1 µl ml^{-1}). 0.5 ml of GES lysis solution (5 M guanidinium thiocyanate (Fisher Bioreagents), 0.5 M EDTA pH 8.0; 0.5% v/v sarkosyl) was added. Following a 15 min incubation at 60 °C, 0.25 ml of cold 7.5 M ammonium acetate (Fisher Bioreagents) were added. After 10 min on ice, 0.5 ml of chloroform (Fisher Bioreagents) were added and samples were vortexed and centrifuged. The aqueous phase was mixed with 0.54 volumes of cold isopropanol (Fisher Bioreagents) and incubated at room temperature for 15 min before centrifugation. The pellet was washed three times in 70% ethanol before resuspending in TE buffer + RNAseA, at which point the concentration was determined spectrophotometrically.

Forward genetic selection for mutants that grow without added CO_2

To select for spontaneous mutants that grow without addition of 5% CO_2 to the atmosphere, wild-type *B. ovis* ATCC 25840 or *B. ovis* Δ*bcaA* strains were inoculated into BB at a final optical density (600 nm; OD_{600}) of ~0.3 then placed in a shaker (200 rpm) under standard atmospheric conditions (0.04% CO_2). Cultures were incubated at 37°C and spectrophotometrically monitored for evidence of growth every 24 hours (**Fig. S1A)**. A total of 29 independent spontaneous mutants were selected that acquired the ability to grow without added CO_2. These mutant isolates were back diluted to OD_{600}=1.5×10^{-3}, 1.5×10^{-5}, 1.5×10^{-7}, 1.5×10^{-9} and grown in 0.04% CO_2 to confirm that the isolates indeed grew without the addition of 5% CO_2 (result from inoculum of 1.5×10^{-5} shown in **Table S1**). Samples were confirmed to *B. ovis* by PCR amplification of a *Brucella*-specific gene (BOV_RS05580).

Whole Genome DNA Sequencing (WGS)

Genomic DNA (gDNA) from the parent *B. ovis* strain, which requires 5% CO_2 to grow, and gDNA from 16 independent spontaneous mutants that grow without added CO_2 (evolved from the parent stock) was purified using the standard guanidinium thiocyanate-based procedure described above. DNA sequencing libraries were prepared from randomly sheared DNA and sequenced using the standard Illumina protocol (Illumina HiSeq 4000; single end 50 bp reads). Reads were mapped to the *Brucella ovis* ATCC 25840 genome (chromosome 1 and chromosome 2 RefSeq accession numbers NC_009505 and NC_009504, respectively) and polymorphisms were identified using breseq (29) (https://github.com/barricklab/breseq). Raw DNA sequencing reads for all strains are available in the NCBI Sequence Read Archive (SRA), through BioProject accession PRJNA540707.

Growth Curves

Strains were struck on SBA or TSBA plates and allowed to grow for ~48 hours. Initial inoculum into BB ranged from OD_{600} of 0.015 to 0.03. Antibiotics for plasmid maintenance and IPTG for expression induction were added when needed at the start of the growth experiment. Growth of two to three separate tubes as technical replicates per sample was monitored spectrophotometrically at 600 nm (OD_{600}). Each growth curve was measured at least three times independently.

Sequence alignments

Genome sequences of *Brucella* isolates were downloaded from PATRIC (https://www.patricbrc.org/) (30). The amino acid sequence was aligned using the multiple alignment tool in Geneious 10.0.9 (www.geneious.com): global alignment with free end gaps, Blosum 62 scoring matrix, gap open penalty of 12 and gap extension penalty of 3. To cluster the *Brucella abortus bcaA* alleles, nucleotide sequences were aligned using Geneious multiple alignment, global alignment with free end gaps specifically, cost matrix set at 65% similarity, gap open penalty of 12, gap extension penalty of 3 and 2 refinement iterations.

RNA sample preparation and RNA sequencing

Samples for RNA sequencing were prepared as follows (see also **Fig. S3A**): six independent cultures each of wild-type *B. ovis* or *B. ovis* $bcaA1_{BOV}$ strain were inoculated into BB and allowed to grow to saturation overnight at 37°C in a roller in the presence of 5% CO_2. Cells were back diluted to an OD_{600} of 0.05 and monitored until they reached OD_{600} of 0.1. Three tubes of *B. ovis* $bcaA1_{BOV}$ were transferred to a roller in a standard air incubator (i.e. 0.04% CO_2), while three tubes of $bcaA1_{BOV}$ remained in the 5% CO_2 incubator. Both sets of $bcaA1_{BOV}$ cultures were harvested after an additional 2.5 hours of growth. Cells from three tubes of wild-type *B. ovis* were harvested at this same time, while the last three tubes of wild-type *B. ovis* were transferred to the standard air incubator and rolled for an additional 2.5 hours. Wild-type *B. ovis* cultures showed no signs of growth over this period in air (0.04% CO_2). Cells from the remaining three wild-type *B. ovis* tubes incubated in the atmospheric (0.04% CO_2) incubator were then harvested.

RNA was prepped from all harvested cell samples as follows: each individual culture sample was aliquoted into six 1.5 ml Eppendorf tubes (for a total of 9 ml) and spun for 60 s at 14,500 × g. Pelleted cells were immediately re-suspended in 1 ml of Trizol (Ambion, Life Technologies, ThermoFisher Scientific) and stored at -80 °C. Samples were then thawed at 65 °C for 10 minutes. Once thawed, 200 µl of cold chloroform was added, cells were vortexed for 15 s and incubated for 5 min at room temperature. Samples were then centrifuged at 4 °C (17000 × g) for 4 min and 500 µl of 100% cold isopropanol was added to the clear supernatant in fresh tube, then frozen for at least 1hr at -80 °C. After 3 min, samples were centrifuged at 4 °C (17000 × g), the supernatant was removed, and 1 ml of 70% ethanol was added to wash the pellet. After an additional 5 min centrifuge (at 4 °C, 17000 × g), residual ethanol was removed and the RNA pellet was resuspended in RNAse free water (50 µl). RNA was treated with DNAse and further purified with RNA purification kit (Qiagen). Concentrations were determined spectrophotometrically, and RNA quality was initially assessed by running 1 µl of each sample on a Tris/Borate/EDTA (TBE) 2% agarose gel. Ribosomal RNA was depleted from the sample using the Gram-Negative bacteria Ribo-Zero rRNA Removal Kit (Illumina-Epicentre). RNA-seq libraries were prepared with an Illumina TruSeq stranded RNA kit according to

Table 1. Sequence polymorphisms[a] between 16 independent *B. ovis* mutants that grow without CO_2 supplementation and the wild-type parent strain[b]

Mutant	Chromosome	Position	Mutation	Gene (or position of intergenic polymorphism)[c]	Gene Annotation	Description
01	NC_009505	805,096	C → T	BOV_RS03980	NADH-quinone oxidoreductase subunit B	T188T (ACC → ACT)
	NC_009505	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
02[d]	NC_009505	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
03	NC_009505	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
04	NC_009505	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
05	NC_009505	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
06	NC_009505	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
07	NC_009505	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
08	NC_009505	1,768,984	-A	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA3*$_{BOV}$ (frameshift)
09	NC_009505	301,445	A → G	BOV_RS0175	IS5/IS1182 family transposase	L224P (CTT → CCT)
	NC_009505	301,958	A → C	BOV_RS01475	sensor histidine kinase	V53G (GTT → GGT)
	NC_009505	420,434	G → T	BOV_RS02035 ← / ← BOV_RS16215	hypothetical protein / hypothetical protein	intergenic
	NC_009505	1,725,552	C → T	BOV_RS08460	branched-chain amino acid ABC transporter ATP-binding protein	G349D (GGC → GAC)
	NC_009505	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
	NC_009504	117,883	(T)8 → (T)7	BOV_RS10890 → / → BOV_RS12585	HTH-type quorum sensing-dependent transcriptional regulator VjbR	intergenic
	NC_009504	676,505	G → A	BOV_RS13635	sulfurtransferase FdhD	pseudogene (415/810 nt)
	NC_009504	908,237	C → T	BOV_RS14715 → / → BOV_RS14720	alpha-hydroxy-acid oxidizing protein lldD / porin family protein	intergenic
10	NC_009505	805,096	C → T	BOV_RS03980	NADH-quinone oxidoreductase subunit B	T188T (ACC → ACT)
	NC_009505	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
11	NC_009505	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
12[d]	NC_009505	1,712,213	G → T	BOV_RS08400	polyprenyl synthetase family protein	A16S (GCC → TCC)
	NC_009505	1,768,979	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA4*$_{BOV}$ (frameshift)
13	NC_009505	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
14	NC_009505	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
15	NC_009505	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
16	NC_009505	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)

[a]All mutations shown in the Table are observed in 100% of the reads from the indicated strain

[b]When comparing the sequence of the *B. ovis* ATCC 25840 strain deposited in GenBank to our ATCC 25840 parent strain, 4 differences were discovered. An insertion of a single G insertion in BOV_RS01360 on chromosome I (NC_009504), at 282,040:1. On chromosome II, at position 20:1, a C insertion at intergenic region and a G insertion at intergenic region in position 459,010:1 These differences were also present in all mutant strains. In ~56% of the sequence reads from the parental strain, we observed a C -> T transition (resulting in a T14I coding change) in BOV_RS10895, *vjbR* at position 118,034 of chromosome II (NC009504). Each derived mutant strain carried one of several different *vjbR* polymorphisms, each fixed at 100% of the sequencing reads.

[c]For intergenic polymorphysims, ← or → indicate direction of flanking genes

[d]This strain lacks the insertion in BOV_RS01360 observed in our parental strain, and thus is identical to the deposited *B. ovis* ATCC 25840 GenBank sequence strain at this locus.

manufacturer's instructions. The libraries were sequenced on an Illumina HiSeq 4000. Sequencing reads were deposited in the NCBI GEO database and are available under accession GSE130678.

RNA sequencing data analysis

Reads were mapped to the *Brucella ovis* ATCC 25840 genome (chromosome I and chromosome II RefSeq accession numbers NC_009505 and NC_009504, respectively) using CLC Genomics Workbench 11.0 (https://www.qiagenbioinformatics.com); mismatch cost: 2; insertion cost: 3; deletion cost: 3; length fraction; 0.8, similarity fraction: 0.8. Wild type *B. ovis* (5% CO_2) replicate 3 and *B. ovis* $bcaA1_{BOV}$ (0.04% CO_2) replicate 1 samples were extreme outliers based on principle component analysis of the raw expression values. This was likely due to RNA sample degradation in these samples during library preparation. Accordingly, these samples were not included in our differential expression analysis. Reads Per Kilobase per Million (RPKM), Transcripts Per Million (TPM), and Counts Per Million (CPM) values for single samples as well as paired differential expressions – with p-values, False Discovery Rate (FDR) and Bonferroni corrections – are presented in **Data Set 1**. Significance thresholds were arbitrarily set at FDR p-value < 0.001 and fold change > |2|. KEGG analysis of significantly regulated genes in wild type *B. ovis* in low (0.04%) versus high (5%) CO_2 conditions was performed using KEGG pathways (31) (https://www.genome.jp/kegg/).

KEGG ontology annotations were assigned to these genes. This annotated gene list was submitted to KEGG Mapper–Search Pathways, querying against *Brucella ovis* ATCC 25840 (bov). Pathways that had 2 or fewer genes were removed. We removed from consideration pathways where the number of upregulated versus downregulated genes was not significantly different. Specifically, pathways where the number of the upregulated genes divided by the total number of genes assigned to that pathway was between 0.4 and 0.6 were removed.

Construction and mapping of *B. ovis* ATCC 25840 Tn-Himar mutant library

B. ovis ATCC 25840 was struck on an SBA plate and incubated for 48h before mating with *E. coli* APA752 (WM3064 donor strain carrying pKMW3 mariner transposon vector library) (32). *E. coli* APA752 was incubated overnight prior to mating (1 ml was thawed and inoculated into 25 ml of LB + DAP + kan). 10 ml of donor strain were pelleted and mixed with pellet of *B. ovis* ATCC 25840 scraped off an SBA plate at a 1:10 ratio, and resuspended in a total of 250 µl of BB. 50 µl of the conjugation mixture were spotted on SBA + DAP plates and left to incubate overnight. Spots were then harvested and resuspended in 5 ml of BB. Cells were diluted to OD_{600} of 0.032 and 500 µl were plated on each of twelve 150 mm SBA + kan plates. Colonies grew in approximately three days and were harvested and resuspended in 250 ml of BB + kan at a final OD_{600} of 0.2. Cells were allowed to double 2-3 times before being frozen in 15% glycerol. For mapping, the library was thawed and the cells were washed once in PBS before pelleting for DNA extraction. The Illumina sequencing library to map Tn-Himar insertion sites was built following protocol of Wetmore *et al*. (32) and as previously described (33, 34). Briefly, adapters were added to the genomic library by PCR (30 min at 95 °C and 15 min at 70 °C) using Mod2_TS_Univ and Mod2_TruSeq primers. DNA was then sheared to obtain fragments approximately 300 to 500 bp in length. Samples were then size selected, end repaired, ligated to a custom adapter, and cleaned before 150 bp single end sequencing by the University of Chicago Functional Genomics Facility. See **Table 3** and **Data Set 5** for list of sequencing statistics and unhit genes, respectively.

Statistical analysis

In bar graphs, error bars represent standard deviation (unless otherwise indicated) of replicates from at least three independent experiments (two or more technical replicates were used for each individual experiment). Strains that failed to grow are indicated as "no growth". **** indicates significance of $p<0.0001$, *** indicates significance of $p<0.001$, ** is $p<0.01$; ns = non-significant, calculated using one-way ANOVA followed by Tukey's test using GraphPad Prism version 8.0.2. For scatter plot, R^2 values were calculated by implementing linear regression in GraphPad Prism. For heat map, z-scores were calculated as follows: for each gene, the CPM value was multiplied by the mean of CPM values for that gene across all samples. The resulting value was divided by the standard deviation of the CPM values of that gene across all samples. Samples were grouped by hierarchical clustering (uncentered correlation; average linkage) using Gene Cluster 3.0 (35, 36) and heat map was built using Java Tree View version 1.1.6r4 (http://jtreeview.sourceforge.net

Results

A forward genetic selection identifies *B. ovis* mutant strains capable of growth in an unsupplemented atmosphere

Brucella ovis requires 5% CO_2 supplementation to grow axenically (25). We sought to identify genes linked to this growth requirement, and thus developed a forward genetic selection to identify spontaneous *B. ovis* mutants that grow in a standard, unsupplemented atmosphere (0.04% CO_2). Specifically, we inoculated wild-type *B. ovis* in Brucella broth (BB) and allowed the cultures to shake in an air incubator. After several days, some cultures became visibly turbid providing evidence for growth (see **Fig. S1A** and **Materials and Methods**). Twenty-nine mutant strains were isolated from five independent selections. We inoculated all mutants into fresh broth and confirmed that they indeed grew without CO_2 supplementation **(Table S1)**. Genomic DNA from 16 of these spontaneous mutants was purified and sequenced along with DNA from the wild-type *B. ovis* parent. Each mutant carried a single

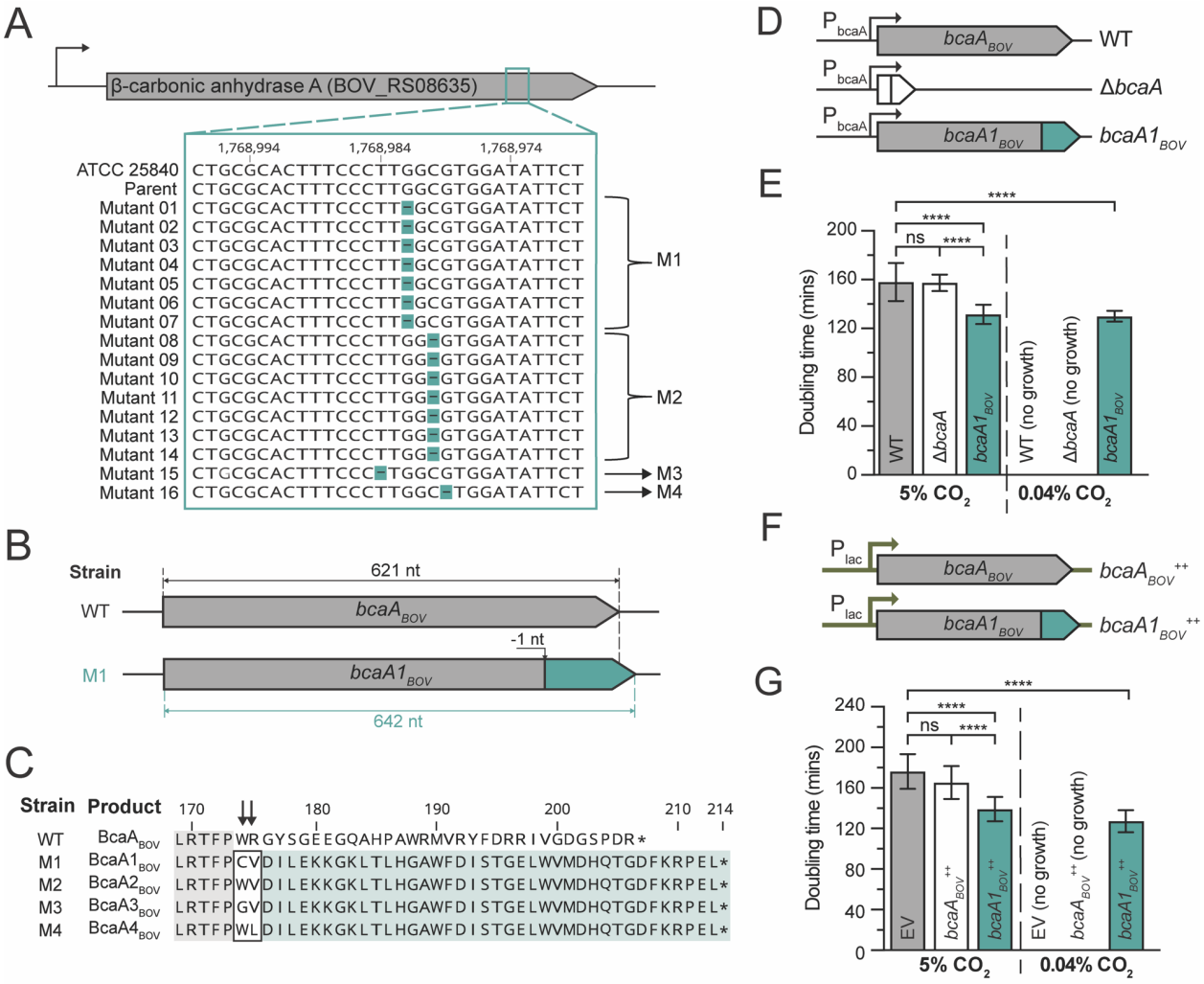


Figure 1. A single nucleotide deletion at the 3' end of *bcaA* enables *B. ovis* growth without CO_2 supplementation
A) Alignment of a segment of the 3' end of *bcaA* (BOV_RS08635) from wild-type *B. ovis* and the 16 selected mutants that can grow without added CO_2. Gaps show the single nucleotide deletions **(highlighted teal dashes)** in the *bcaA* locus. M1, M2, M3 and M4 are the four classes of single nucleotide deletions that we observed. Annotated nucleotide position is indicated at the top. **B)** Schematic representation of the wild-type *B. ovis* ATCC 25840 allele, $bcaA_{BOV}$ **(top)**, compared to the frameshifted mutant allele from the M1 cluster of mutants, $bcaA1_{BOV}$ **(bottom)**. The site of the frameshift is indicated with a vertical arrow. **Teal** shading indicates the portion of the gene that has an altered coding sequence following the frameshift that results from deletion of a single guanosine at position 1,768,986. **C)** Multiple protein sequence alignment of the C-terminal frameshift site of wild-type *B. ovis* ATCC 25840 $BcaA_{BOV}$ and the four selected mutant alleles (M1 to M4). **Teal** shaded area highlights the frameshifted amino acids. **Gray** shaded area highlights sequence that is conserved in all five alleles. **Boxed** amino acids show protein sequence variation at the frameshift site **(black arrows)**. **D)** Schematic of the *bcaA* locus in strains used for experiments in panel **E**; wild-type *B. ovis* ATCC 25480 $bcaA_{BOV}$ allele **(top)**, in-frame $bcaA_{BOV}$ deletion allele (Δ*bcaA*) **(middle)**, and a *B. ovis* strain in which the wild-type allele is replaced with the frameshifted $bcaA1_{BOV}$ allele at the native locus **(bottom)**. P_{bcaA} indicates expression from the native *bcaA* promoter. **E)** Doubling times in hours for the three strains in **panel D** grown with either 5% CO_2 supplementation or in air without added CO_2 (0.04%). Error bars represent standard deviation of replicates from three independent experiments (each performed with at least two technical replicates), for a total of 8 measurements per sample. Strains that failed to grow are indicated with "no growth". **** indicates significance of $p<0.0001$, calculated using one-way ANOVA followed by Tukey's post-test. **F)** Schematic of RK2-derived plasmids (pSRK) harboring two different *bcaA* alleles. pSRK carrying a *lac* inducible (P_{lac}) wild-type $bcaA_{BOV}$ **(top)** or frameshifted $bcaA1_{BOV}$ **(bottom)** were transformed into wild-type *B. ovis* ATCC 25840. Plasmid-bearing strains are referred to as $bcaA_{BOV}^{++}$ and $bcaA1_{BOV}^{++}$, respectively. **G)** Doubling times (in hours) of strains carrying pSRK-$bcaA_{BOV}$ ($bcaA_{BOV}^{++}$), pSRK-$bcaA1_{BOV}$ ($bcaA1_{BOV}^{++}$) or the empty vector (EV) as control. Cells were induced with 1mM IPTG, grown in BB with kanamycin to maintain plasmids, and cultivated in a standard air incubator (0.04% CO_2) or with 5% CO_2 supplementation. Error bars represent standard deviation of four independent experiments executed in triplicate (total number of measurements per sample = 12). Strains that failed to grow are indicated with "no growth". *** indicates significance of $p<0.0001$, calculated using one-way ANOVA followed by Tukey's post-test.

nucleotide deletion within a six-nucleotide interval at the 3' end of a β-carbonic anhydrase gene (locus tag BOV_RS08635; $bcaA_{BOV}$) (**Fig. 1A and Table 1**). These deletions in $bcaA_{BOV}$ frameshifted the coding sequence of the $BcaA_{BOV}$ C-terminus and increased the length of the predicted gene product from 207 to

214 residues (**Fig. 1A-C**). The 16 sequenced mutants clustered into 4 groups based on the site of the frameshift mutation; the corresponding alleles are named *bcaA1*$_{BOV}$-*4*$_{BOV}$ (**Fig. 1C**). Nucleotide deletions in *bcaA*$_{BOV}$ were the only mutations that were common to all mutant strains and missing from the parent (**Table 1**).

Similar *bcaA*$_{BOV}$ mutations in *B. ovis* strains that grow without CO_2 have been recently described by Perez-Etayo *et al*.(27), though they have used the name carbonic anhydrase II (CAII) for the gene locus and product. For the purposes of this manuscript, we have chosen to use the bacterial genetic naming guidelines of Demerec *et al*. (37), and thus refer to this gene as *bcaA*$_{BOV}$. We conclude that cultivation of wild-type *B. ovis* in a standard atmosphere selects for a frameshift mutation at the 3' end of the β-carbonic anhydrase gene, *bcaA*$_{BOV}$. This frameshift increases the length of the predicted protein and completely changes the sequence of the last 32-33 residues of BcaA$_{BOV}$. Based on the fact that we successfully selected mutants in air after several days of incubation, we presume that wild-type *B. ovis* can replicate (albeit very slowly) in an unsupplemented atmosphere.

A *bcaA* frameshift enables growth in an unsupplemented atmosphere

To directly test the hypothesis that *bcaA*$_{BOV}$ frameshift mutations enable growth in an unsupplemented atmosphere, we built a strain in which wild-type *bcaA*$_{BOV}$ was replaced with the frameshifted *bcaA1*$_{BOV}$ allele. We also built a strain in which *bcaA*$_{BOV}$ was deleted in-frame (Δ*bcaA*) (**Fig. 1D)**. We measured growth of these strains and wild-type *B. ovis* in 5% CO_2, and in standard atmospheric CO_2 levels (approx. 0.04%). All three strains grew in the presence of 5% CO_2, indicating that *bcaA*$_{BOV}$ is not required for growth in a CO_2-supplemented atmosphere. When these strains were incubated with 0.04% CO_2, only the *bcaA1*$_{BOV}$ allele replacement strain showed measurable growth. Notably, the *bcaA1*$_{BOV}$ strain grew significantly faster (i.e. had a shorter doubling time) than wild-type and Δ*bcaA* strains in the presence of 5% CO_2 (**Fig. 1E and Fig. S1B**). From these data, we conclude that frameshift mutations at the 3' end of *bcaA1*$_{BOV}$ enable robust growth of *B. ovis* in an unsupplemented atmosphere and result in faster growth than wild-type in the presence of 5% CO_2.

Table 2. *B. ovis* ATCC 25840 Tn-Himar library

Estimated number of Tn strains[a]	1.8×10^6
Unique Tn insertions passing mapping criteria	83,660
Numbers of unique TA sites hit[b]	50,984
Unique insertions in central 10%-90% of genes	52,612
Median Tn per protein-coding gene	15
Number of protein-coding genes hit	2,753 (82.7%)
Number of unhit genes[c]	488 (14.4%)

[a]The number represents Chao2 estimate of the total number of barcodes in the library
[b]Sites on opposite strand count as unique
[c]Genes that are more than 300nt with good insertions were counted as 'unhit'

***bcaA* frameshift mutations are the sole genetic mechanism by which *B. ovis* growth is spontaneously rescued under low CO_2 tension**

To test whether mutations in genes other than *bcaA*$_{BOV}$ enable *B. ovis* growth in an unsupplemented atmosphere, we repeated the forward genetic selection described above starting with the Δ*bcaA* strain. We incubated Δ*bcaA* in BB at 37°C for over two weeks in an air incubator and never observed turbid cultures. This experiment was repeated 4 times, with approximately 30 inoculated tubes each time, and growth was never observed. This stands in contrast with our initial selection experiment, in which multiple tubes inoculated with wild-type *B. ovis* became turbid over this timescale. As an alternative forward genetic approach, we generated a pool of *B. ovis* Tn-Himar mutants with insertions at over 5×10^4 unique sites (**Table 2)**. This experiment was aimed at identifying insertional mutations that enable growth in a standard atmosphere. We incubated aliquots of this pool in BB for over two days in an air incubator and did not observe growth. Thus, none of the transposon insertions in the pool enabled *B. ovis* growth in the absence of CO_2 on this short timescale. From these data, we conclude that single nucleotide deletions at the 3' end of *bcaA*$_{BOV}$, which result in a frameshifted gene product, are the sole genetic mechanism by which *B. ovis* spontaneously acquires the ability to grow without added CO_2 on this experimental timescale.

The frameshifted alleles, *bcaA1*$_{BOV}$-*4*$_{BOV}$, are dominant over the wild-type *bcaA*$_{bov}$ allele

We next transformed wild-type *B. ovis* with low-copy replicating plasmids, from which we expressed either the wild-type (*bcaA*$_{BOV}^{++}$) or frameshifted (*bcaA1*$_{BOV}^{++}$) alleles from an IPTG-inducible (*lac*) promoter (P_{lac}) (**Fig. 1F**). A strain carrying an empty vector (EV) was used as a control. The EV and the *bcaA*$_{BOV}^{++}$ strains (with 1 mM IPTG) grew comparably in 5% CO_2 and did not grow when shifted to a standard atmosphere (0.04% CO_2) (**Fig. 1G and Fig. S1C**). Thus, overexpressing wild-type *bcaA*$_{BOV}$ from a plasmid is not sufficient to enable growth in a standard atmosphere.

The *bcaA1*$_{BOV}^{++}$ strain grew without added CO_2, demonstrating that the frameshifted allele is dominant and sufficient to enable *B. ovis* growth in 0.04% CO_2.

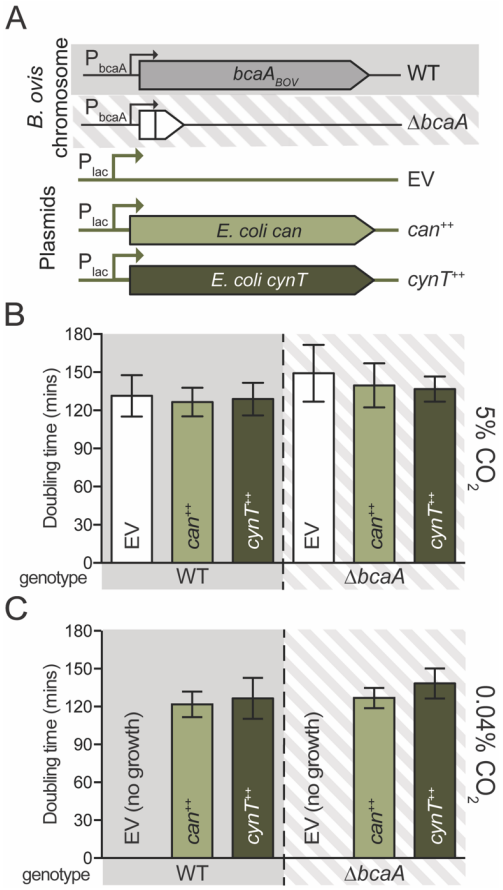


Figure 2. Heterologous expression of two different *Escherichia coli* β-carbonic anhydrases enables growth of wild-type *B. ovis* ATCC 25840 without CO_2 supplementation

A) Schematic representation of RK2-derived plasmids (pSRK) transformed into either wild-type *B. ovis* ATCC 25840 (WT; **gray background**) or the in-frame *bcaA* deletion strain (Δ*bcaA*, **striped gray background**). pSRK plasmids carried either *E. coli* β-carbonic anhydrase *can* (*can*$^{++}$) or *cynT* (*cynT*$^{++}$) under inducible control of P_{lac}. Empty vector (EV) plasmid was used as a control (**green line**). **B and C)** Doubling time (in minutes) of strains outlined in **A** grown in 5% CO_2 **(B)** or 0.04% CO_2 **(C)** after induction with 1mM IPTG. Cells were grown in BB with kanamycin to maintain plasmids. The parent genotype (wild type or Δ*bcaA*) is indicated on the bottom of each bar graph. The plasmid genotypes are indicated in each bar. Experiment was performed in triplicate two to four independent times (with two to three technical replicates for a total of 5 to 14 measurements per sample). Error bars represent standard deviation of at least two independent experiments. Strains that failed to grow are indicated with "no growth".

Moreover, the strain expressing *bcaA1*$_{BOV}$$^{++}$ grew significantly faster in 5% CO_2 than the *bcaA*$_{BOV}$$^{++}$ and EV strains, providing additional evidence that *bcaA1*$_{BOV}$ enhances growth rate in broth culture (**Fig. 1G and Fig. S1C**).

To confirm that all four classes of frameshifted *bcaA* alleles identified in our forward genetic selection (*bcaA1*$_{BOV}$-*4*$_{BOV}$) were sufficient to enable growth in a standard atmosphere, we expressed these alleles from a replicating plasmid in a standard air incubator (0.04% CO_2) and compared growth to the *B. ovis* EV strain. As expected, all four alleles enabled growth in 0.04% CO_2 (**Fig. S1D**). Thus, different single nucleotide deletions at the 3' end of *bcaA*$_{BOV}$ (**Fig. 1C**) result in the same phenotype: growth of *B. ovis* in a standard atmosphere. Given that the four alleles phenocopy each other, we continued our study with the *bcaA1*$_{BOV}$ allele. This was the most common isolate from our forward genetic selection and is the most conserved *bcaA* allele in the entire *Brucella* clade (see below).

Expression of *E. coli* β-carbonic anhydrases is sufficient to enable *B. ovis* growth in a standard atmosphere

bcaA$_{BOV}$ is predicted to encode a β-class carbonic anhydrase. These enzymes catalyze the reversible hydration of CO_2 to bicarbonate, an important anabolic substrate. Carbonic anhydrases have been extensively studied (28, 38) in diverse species (39-41), including *Brucella suis* 1330 (42-44). The thirty-three C-terminal residues of wild-type *B. ovis* BcaA$_{BOV}$ differ from *B. suis* 1330 BcaA. However, the sequence of the frameshifted mutant *B. ovis* BcaA1$_{BOV}$ protein is identical to *B. suis* BcaA$_{BSU1330}$ (locus tag: BR_RS08400; also known as $_{Bs1330}$CAII (27) or bsCAII (42)) with the exception of one residue (**Fig. S2A**). The *B. ovis bcaA1*$_{BOV}$ mutation thus results in a frameshifted gene product with a C-terminus that matches *B. suis* BcaA$_{BOV}$ and other *Brucella* species, which we discuss below.

B. suis 1330 BcaA$_{BSU1330}$ is reported to be an active β-carbonic anhydrase (42, 43). Thus, we reasoned that the selected frameshift mutations at the 3' end of *B. ovis bcaA*$_{BOV}$ restored the protein to an active form that can hydrate CO_2 at the low partial CO_2 pressure of a standard atmosphere, thereby enabling growth. Following this hypothesis, we tested whether heterologous expression of known active *Escherichia coli* MG1655 β-class carbonic anhydrases *can* (locus tag: b0126) or *cynT* (locus tag: b0339) (45) was sufficient to support growth of the wild-type and Δ*bcaA B. ovis* strains in a standard atmosphere **(Figure 2A)**. Induction *can* or *cynT* expression from a low-copy plasmid (RK2-P_{lac}) with 1mM IPTG had no effect on growth rate in 5% CO_2 (**Fig. 2B**), but enabled growth of wild-type and Δ*bcaA B. ovis* in an unsupplemented atmosphere (0.04% CO_2) (**Fig. 2C**). We conclude that expression of *E. coli can* or *cynT* (**Fig. S2B**) is sufficient to enable growth of WT and Δ*bcaA B. ovis* strains under the low partial CO_2 pressure of a standard atmosphere.

A comparative analysis of *bcaA* sequences and the CO_2 growth requirement in the genus *Brucella*

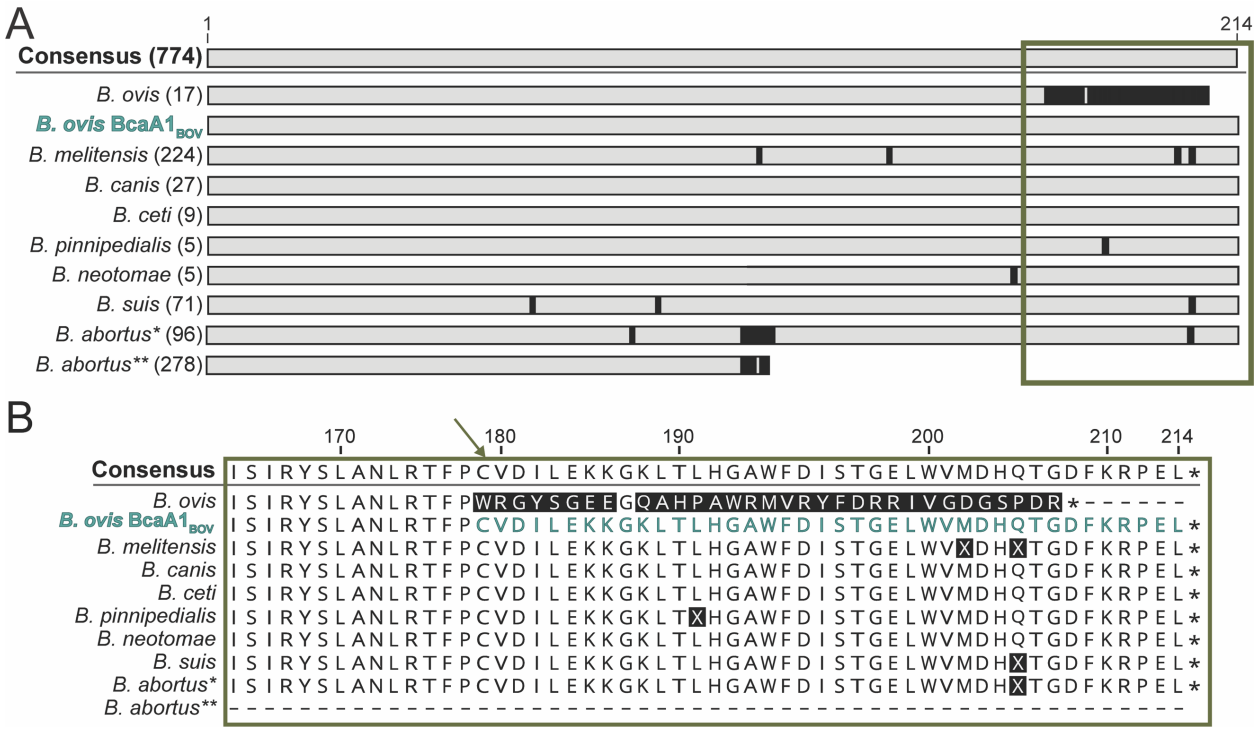
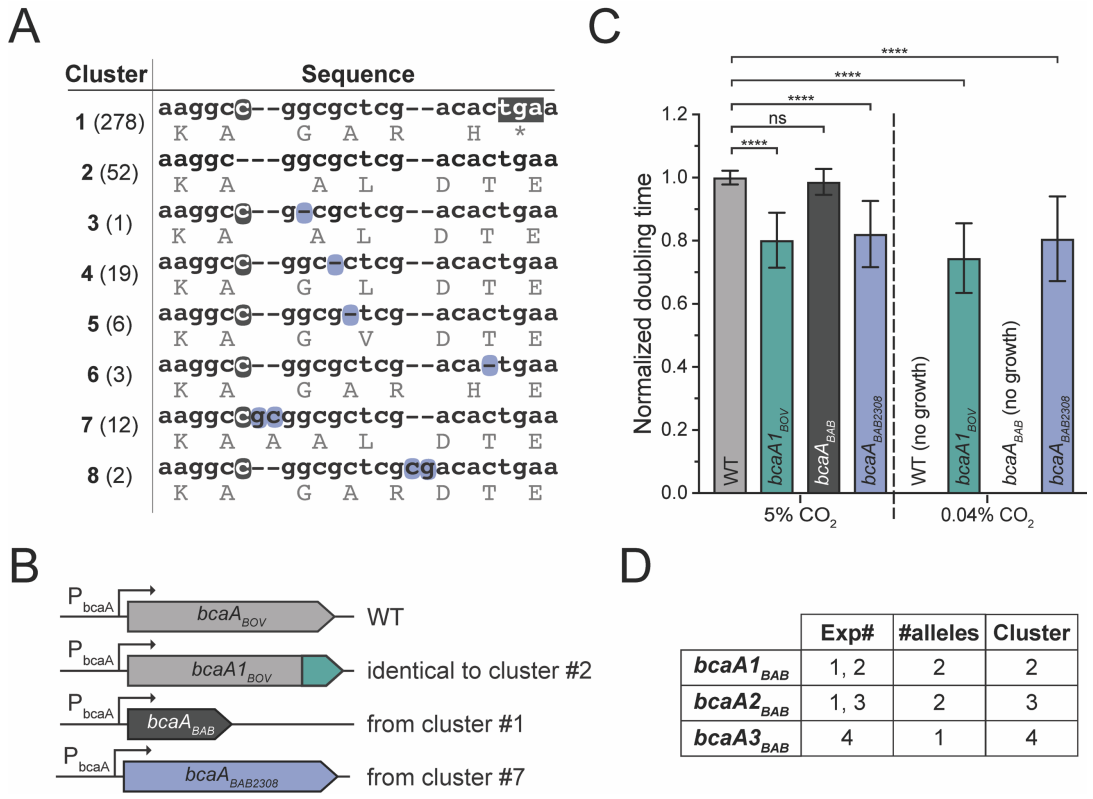


Figure 3. Sequence polymorphisms in BcaA orthologs across the genus *Brucella*
A) Schematic of the BcaA protein sequence polymorphisms based on alignment of 774 BcaA orthologs from *Brucella* species available in the PATRIC database and found in **Data Set 1**, including *B. ovis* BcaA1$_{BOV}$ (**teal**; M1 cluster). Genus level consensus of all sequences is depicted on top. Species-specific consensus sequences were generated for species with at least 5 representatives (sequences not included in **Fig.3** are grouped in 'Other, **Data set 1**). The species-specific consensuses were then aligned as shown. Amino acid differences between the species-specific and genus-level BcaA consensus sequences are indicated in **black**. *Brucella abortus* sequences were split into two groups, *B. abortus** and *B. abortus***, based on the presence or absence of a frameshift (and premature stop codon) resulting from a single cytosine insertion 339 nucleotides after the annotated start codon, respectively (**see also Fig. 4 and main text**). The number of strains compiled in each species-specific consensus sequence is shown in parentheses. The frameshifted BcaA1$_{BOV}$ allele that enables growth of *B. ovis* without added CO_2 is included in the alignment (**teal**). **B**) Enlargement of green boxed area in (**A**). The C-terminal end of *B. ovis* BcaA1$_{BOV}$ (M1 cluster) is highlighted in **teal**; the site of the frameshift is indicated (**green angled arrow**).

As outlined in the introduction, *B. ovis* and the majority of *B. abortus* strains require CO_2 supplementation for axenic cultivation (26). To assess the extent to which variation in *bcaA* is linked to this physiologic trait, we compared the sequences of 773 *bcaA* orthologs from the genus *Brucella* (see **Data Set 1**). *B. ovis* and the majority of *B. abortus* encode BcaA variants that significantly deviate from the consensus sequence for the genus (**Fig. 3A**).

All 17 *B. ovis* $bcaA_{BOV}$ sequences in our dataset harbor an additional guanosine 522 nucleotides after the start codon (G522) relative to other *Brucella* species **(Fig. 1)**, which results in a corresponding frameshifted BcaA$_{BOV}$ C-terminus (**Figs. 1 & 3**). The *B. ovis* sequences we analyzed were from strains isolated from sheep in Europe, South America, Australia, and New Zealand over the course of several decades (**Data Set 1**). Thus, this single nucleotide insertion in $bcaA_{BOV}$ is apparently a defining genetic characteristic of *B. ovis*. As highlighted in **Figs. 1A-C and 3B,** single nucleotide deletion mutations that enable *B. ovis* growth in a standard atmosphere ($bcaA1_{BOV}$-4_{BOV}) abolish this frameshift and restore the BcaA C-terminus to the consensus of the *Brucella* genus.

We further analyzed *bcaA* from 374 *B. abortus* strains and observed two major sequence classes: 1) the majority of sequences (278) encode a *bcaA* pseudogene (*B. abortus*** in **Fig. 3**), and 2) the remainder encode full-length proteins (*B. abortus** in **Fig. 3**) that completely or almost completely match consensus (see below)**.** The *bcaA* pseudogene ($bcaA_{BAB}$) sequences all share the same C339 insertion that results in a frameshift and premature stop codon (**Fig. 4A, cluster 1**). The *bcaA* sequences from the remaining 96 *B. abortus* strains, encoding a full-length BcaA protein, grouped into seven clusters. The largest of these clusters comprised sequences lacking the C339 insertion (52 out of 374 strains), resulting in a consensus BcaA gene product **(Figs. 3 and 4A, cluster 2)**. In the other *B. abortus bcaA* clusters the consensus reading frame was restored by either a single nucleotide deletion or two-nucleotide insertions within 12 base pairs (downstream) of the C339



	Exp#	#alleles	Cluster
bcaA1$_{BAB}$	1, 2	2	2
bcaA2$_{BAB}$	1, 3	2	3
bcaA3$_{BAB}$	4	1	4

Figure 4. Analysis of sequence and function of *B. abortus bcaA* alleles in *B. ovis* ATCC 25840
A) Nucleotide alignment of *B. abortus bcaA* from 374 sequenced isolates starting at position +339 (codon 112). *B. abortus bcaA* sequences cluster into 8 groups based on nucleotide variation. Cluster number is indicated on the left and the number of sequences in each cluster is shown in parenthesis. Translated sequence for each codon is shown in light gray. Cytosine insertion that frameshifts the genus-level consensus (cluster 2) is highlighted **gray**. **Gray box** in cluster 1 (i.e. allele $bcaA_{BAB}$) highlights the premature stop codon arising from C insertion. Second-site, single-nucleotide deletions or double nucleotide insertions (**highlighted in blue**) that restore reading frame to consensus are highlighted for clusters 3 through 8. Cluster 7 represents the *bcaA* allele present in *Brucella abortus* strain 2308 ($bcaA_{BAB2308}$). Clusters 2 and 3 have the same amino acid sequence, but different nt sequence. **B)** Schematic of *bcaA* locus in *B. ovis* ATCC 25840 strains that were functionally assayed in **C**. **From top**: wild-type $bcaA_{BOV}$ allele, the gain-of-function $bcaA1_{BOV}$ allele, the $bcaA_{BAB}$ pseudogene (cluster 1), or the $bcaA_{BAB2308}$ allele (cluster 7), each expressed from the native promoter (P_{bcaA}). **C)** Normalized *B. ovis* doubling times comparing WT ($bcaA_{BOV}$ allele) to strains harboring $bcaA1_{BOV}$, $bcaA_{BAB}$, or $bcaA_{BAB2308}$ in either high (5%) or low (0.04%) CO_2. Error bars represent standard deviation of replicates from three to eight independent experiments for a total of 11 to 23 measurements per strain. **** indicates significance of $p<0.0001$, ns = non-significant, calculated using one-way ANOVA followed by Tukey's post-test. Strains that failed to grow are indicated as 'no growth'. **D)** Tabulation of gain-of-function $bcaA_{BAB}$ mutants that arose in 4 independent selection experiments. Table shows experiment number (Exp#), number of times a specific allele ($bcaA1_{BAB}$-3_{BAB}) was isolated across the four independent experiments (#alleles), and cluster assignment for that allele (Cluster). Cluster designation as per panel **A.**

frameshift site **(Fig. 4A)**. We posit that the $bcaA_{BAB}$ pseudogene is the ancestral allele in *B. abortus*, and that insertion/deletion (i.e. indel) mutations near nucleotide 339 have been selected in different *B. abortus* strains/biovars to resurrect functional BcaA variants that enable growth under low partial CO_2 pressure.

Functional *B. abortus bcaA* alleles are rapidly selected from a $bcaA_{BAB}$ pseudogene under CO_2 limitation

We next assayed the function of two *B. abortus bcaA* alleles ($bcaA_{BAB}$ and $bcaA_{BAB2308}$) in *B. ovis* (**Fig. 4B**). Specifically, we replaced wild-type *B. ovis* $bcaA_{BOV}$ with either the $bcaA_{BAB}$ pseudogene or the near-consensus $bcaA_{BAB2308}$ allele (cluster 7, **Figure 4A**), and measured growth in a standard (0.04% CO_2) unsupplemented atmosphere. Replacing $bcaA_{BOV}$ with $bcaA_{BAB2308}$ enabled growth of *B. ovis* in a standard atmosphere (**Fig. 4C**), consistent with Perez-Etayo *et al.* (27). Replacing $bcaA_{BOV}$ with the frameshifted $bcaA_{BAB}$ pseudogene did not enable *B. ovis* growth in a standard atmosphere (**Fig. 4C**). We conclude that $bcaA_{BAB}$ is inactive *in vivo* and that the majority of *Brucella abortus* isolates cannot grow in an unsupplemented atmosphere, which is consistent with historical observations outlined in our introduction (8, 10, 11, 22). Conversely, approximately one-quarter of sequenced *B. abortus* strains harbor a copy of *bcaA* in which the frameshift

Table 3. *B. abortus* biovar association with *bcaA* allele clusters

Biovar	CO_2 dependence[a]	Total	Cluster 1 (*bcaA*$_{BAB}$)	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7[b] (*bcaA*$_{BAB2308}$)	Cluster 8
Biovar 1[c]	+(–)	196[a] (52%)	179 (89%)	8 (4%)	0	2 (1%)	0	0	7 (3%)	0
Biovar 2	+	18[a] (5%)	13 (72%)	1 (5%)	0	2 (11%)	1 (5%)	1 (5%)	0	0
Biovar 3	+(–)	14 (4%)	10 (71%)	1 (7%)	2 (14%)	1 (7%)	0	0	0	0
Biovar 4	+(–)	16 (4%)	15 (94%)	0	0	0	1 (6%)	0	0	0
Biovar 5	–	3 (1%)	0	0	0	3 (100%)	0	0	0	0
Biovar 6	–	11 (3%)	0	10 (91%)	1 (9%)	0	0	0	0	0
Biovar 7	–	5 (1%)	1 (20%)	4 (80%)	0	0	0	0	0	0
Biovar 9	–or+	6 (2%)	2 (33%)	4 (67%)	0	0	0	0	0	0
Unknown		105[b] (28%)	57 (55%)	23 (22%)	0	11 (10%)	4 (4%)	2 (2%)	5 (5%)	2 (2%)

[a]Symbols indicating CO_2 dependence as per Alton *et al.*(26): + = CO_2 dependent; +(–) = mostly CO_2 dependent; – = CO_2 independent.
[b]Two strains are annotated as 2308 thus possibly biovar 1 and redundant. Biovar was not assigned.
[c]One strain is annotated as "atypical"

mutation at C339 has apparently been remediated **(Fig. 4A)** to restore expression of a full-length protein. Given these results, we hypothesized that functional variants of *B. abortus bcaA*$_{BAB}$ could be experimentally selected in *Brucella* cells cultivated under standard atmospheric conditions, i.e. low pCO_2.

To test this hypothesis, we used the *B. ovis* strain in which we replaced *bcaA*$_{BOV}$ with the non-functional *B. abortus bcaA*$_{BAB}$ pseudogene (carrying the C339 insertion) and selected for mutants that could grow in an unsupplemented atmosphere (**Fig. 4D**). Four independent experiments yielded spontaneous mutants that grew to high density without CO_2 supplementation within several days. We sequenced the *bcaA* locus of 5 independent mutants from these selections and identified three types of mutations (*bcaA1*$_{BAB}$ -*3*$_{BAB}$) that remediated the C339 frameshift and restored the full-length, consensus BcaA coding frame. All *bcaA* variants that we selected experimentally are present in sequenced isolates of *B. abortus*. We cataloged information of the 374 sequenced *B. abortus* isolates, including assigned biovars (when available) from PATRIC (**Table 3**). *B. abortus* biovars 5 and 6 have been described as CO_2 independent (26) and, as predicted, all sequenced isolates of these biovars possess consensus (i.e. full-length) *bcaA* alleles. Biovars 1, 3 and 4 are described as primarily CO_2 dependent, which is consistent with the fact that between 70 and 95% of sequenced isolates from these biovars carry the *bcaA*$_{BAB}$ pseudogene. *B. abortus* biovar 2 is categorized as CO_2 dependent (26), yet 5 of 18 sequenced isolates possess a consensus functional *bcaA* allele; biovar 7 is classified as CO_2 independent, though one of five sequenced isolates carries the *bcaA*$_{BAB}$ pseudogene. We conclude that the identity of *bcaA* alleles in the different *B. abortus* biovars is largely consistent with the CO_2 dependence/independence cataloged by Alton *et al.* (26).

Our analysis shows that the most common allele of *bcaA* in *B. abortus* is the frameshifted *bcaA*$_{BAB}$ pseudogene. Thus, like *B. ovis*, many wild *B. abortus* populations apparently harbor a non-functional, frameshifted *bcaA* pseudogene that renders the bacterium unable to grow without CO_2 supplementation. Also, like *B. ovis*, functional (consensus) *bcaA* alleles can be easily selected from *bcaA*$_{BAB}$ by incubating the bacterium in a standard atmosphere for several days.

CO_2 limitation triggers a large-scale transcriptional response in wild-type *B. ovis*, while transcription in the *bcaA1*$_{BOV}$ strain is insensitive to CO_2 shifts

Our group and others (27) have shown that *bcaA* is a specific genetic determinant of growth under low pCO_2. We have further shown that *B. ovis* and most *B. abortus* strains harbor non-functional *bcaA* pseudogenes containing distinct single-base insertion mutations that result in a frameshifted gene product. These insertional mutations underlie the long-noted CO_2 requirement for axenic cultivation of these species (see **Introduction**). Given these results, we sought to better understand the physiological consequences of CO_2 limitation in *B. ovis,* a species that requires CO_2 for axenic cultivation. Specifically, we defined transcriptional changes in *B. ovis* ATCC 25840 at the genome scale in response a downshift in pCO_2.

To measure the transcriptional response to CO_2 change, we cultivated *B. ovis* wild-type and *bcaA1*$_{BOV}$ strains in a 5% CO_2 atmosphere. A subset of these cultures was then shifted to a standard atmosphere (0.04% CO_2). After a controlled incubation period in this condition (see **Fig. S3A** and **Materials and Methods**) we harvested cells, extracted RNA from all samples, and prepared libraries for sequencing. Shifting wild-type *B. ovis* from 5% to 0.04% CO_2 slowed growth (**Fig. 5A**), and

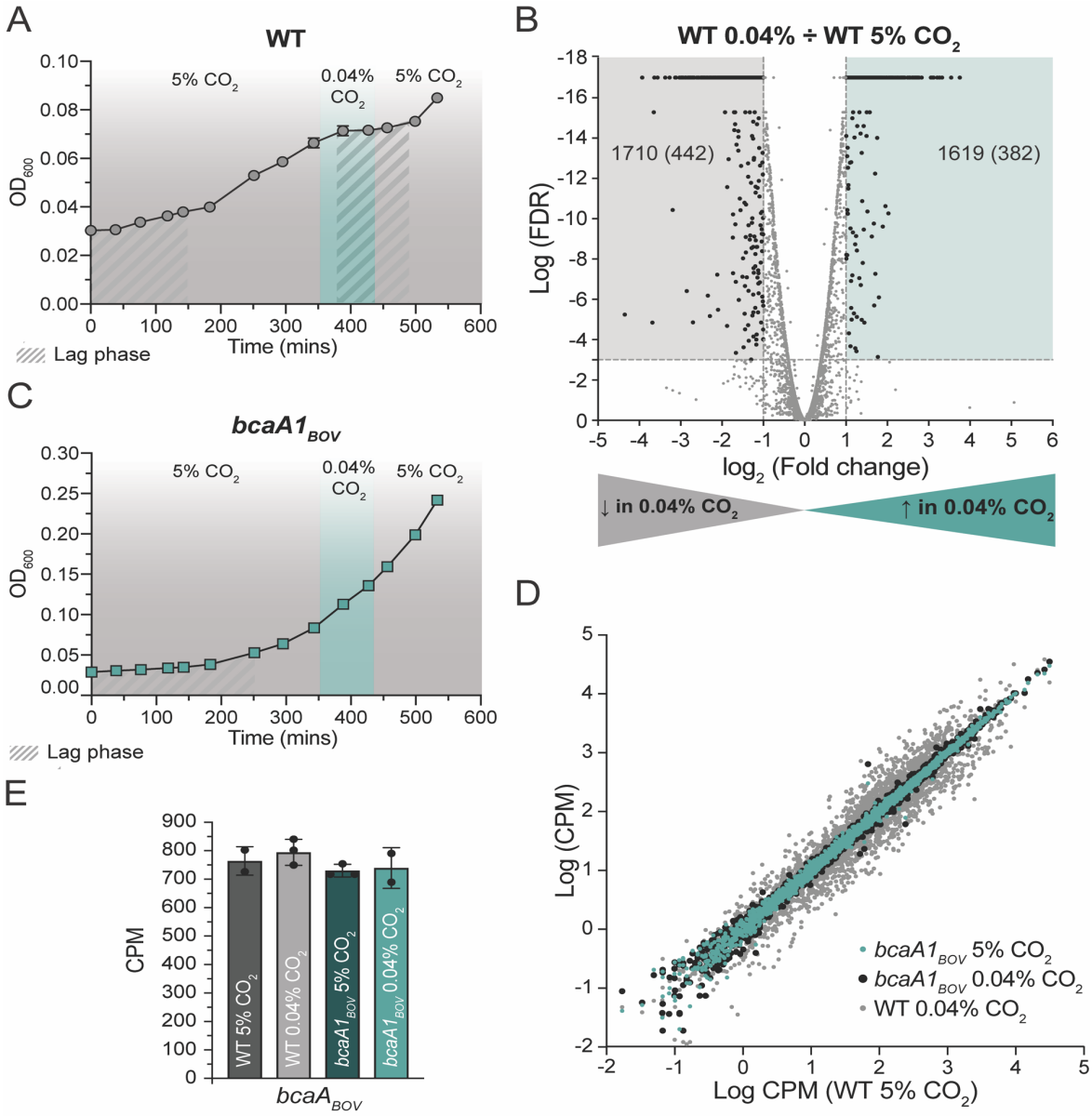


Figure 5. Gene expression changes in wild-type *B. ovis* and *B. ovis bcaA1$_{BOV}$* upon CO_2 downshift from 5% to 0.04%
A) Growth of wild-type *Brucella ovis* ATCC 25840 inoculated in BB and grown in 5% CO_2. In log phase, cells were shifted to 0.04% CO_2 (**teal shading**) for approximately 90 minutes then shifted back to 5% CO_2. Lag phase is highlighted with **gray striped shading**. Experiment was performed three independent times and a representative experiment is shown. **B)** Volcano plot showing $\log_2$ fold change in gene expression versus $\log_{10}$ of the false discovery rate (FDR)-corrected p-value after downshift of wild-type cells from 5% to 0.04% CO_2. Each dot represents a gene. Genes with FDR p-values < 0.001 and with $\log_2$ fold changes > |1| were considered statistically significant and are indicated as **black dots**. Due to floating point value constraints in calculation of p-values, the smallest possible FDR p-value we were able to calculate is 1×10^{-15}. All lower FDR p-values were arbitrarily assigned a value of 1×10^{-17}. Genes that were considered upregulated upon CO_2 downshift are indicated on the right (**teal shading**) and genes that are downregulated upon downshift are indicated on the left (**gray shading**); statistically significant regulated genes in each set are in parenthesis. **C)** Same growth experiment in **(A)** but with the *B. ovis bcaA1$_{BOV}$* strain, which was selected to grow without added CO_2. **D)** Absolute expression of all measurable transcripts, in $\log_{10}$ counts per million (CPM), from wild-type *B. ovis* ATCC 25840 in high (5%) CO_2 compared to *B. ovis bcaA1$_{BOV}$* in 5% CO_2 (**teal**), *B. ovis bcaA1$_{BOV}$* in low (0.04%) CO_2 **(black)** or wild-type in low (0.04%) CO_2 (**gray**). Each dot represents the mean CPM for a gene (see **Data Set 1**). R^2 values were calculated (see **main text** and **Fig. S3C**) implementing linear regression using GraphPad Prism 8. CPMs were calculated using CLC Genomics Workbench 11.0. **E)** Steady-state *bcaA* transcript levels by strain and condition as indicated. Error bars represent standard deviation calculated from two to three independent samples subjected to RNA-seq measurements

induced large-scale changes in gene expression: 382 genes were significantly upregulated and 442 genes were significantly downregulated upon a downshift in pCO_2 ($\log_2$ fold change > |1|; false-discovery rate (FDR) p-value ≤ 0.001) **(Data Set 2 and Fig. 5B)**. In contrast, shifting a *B. ovis* strain harboring the *bcaA1$_{BOV}$* allele from 5% to 0.04% CO_2 had no effect on growth **(Fig. 5C)** or gene expression using the same

significance threshold **(Data Set 2 and Fig. S3B)**. In fact, a comparison of wild-type *B. ovis* cultivated at 5% CO_2 to *bcaA1*$_{BOV}$ cultivated in 0.04% CO_2 revealed no differentially expressed genes. As expected from these results, the overall transcriptional profile of wild-type *B. ovis* cultivated in 5% CO_2 is highly correlated with the *B. ovis* *bcaA1*$_{BOV}$ strain cultivated under either 5% CO_2 or 0.04% CO_2 (R^2= 0.995, 5% CO_2 and R^2= 0.993, 0.04% CO_2) **(Fig. 5D and Fig. S3C)**. We conclude that wild-type *B. ovis* is acutely sensitive to downshifts in levels of atmospheric CO_2, and that presence of *bcaA1*$_{BOV}$ – the consensus allele in the *Brucella* clade – renders *B. ovis* insensitive to CO_2 downshifts. Notably, our data show that transcription of *bcaA*$_{BOV}$ itself is not induced by CO_2 limitation (**Fig. 5E**).

CO_2 limitation induces a stringent-like starvation response and activates transcription of the *virB* type IV secretion gene cluster in wild-type *B. ovis*

We next assigned each of the genes up- or down-regulated upon pCO_2 downshift in wild-type *B. ovis* to Kyoto Encyclopedia of Genes and Genomes (KEGG) categories. Of the 824 genes that had a $\log_2$ fold change > |1| and an FDR p-value < 0.001, 447 could be classified into one or more KEGG categories (**Data Set 3)**. Only categories where 3 or more genes were assigned were considered in our pathway analysis (**Fig. 6)**. Genes involved in translation, including those encoding ribosomal proteins, ribosomal RNA, transfer RNA, and aminoacyl-tRNA synthetases, are largely downregulated upon shift from 5% to 0.04% CO_2 (**Data Set 2, 3 and Fig. 6**). In addition, transcript levels for elongation factors P, Tu and G are highly reduced upon CO_2 downshift. This transcriptional profile, which reflects reduced protein synthesis, is consistent with nutrient limitation and activation of the stringent response (46-49). In addition, multiple respiratory, energy metabolism, transport, and catabolic genes are downregulated upon CO_2 downshift **(Data Set 2, 3 and Fig. 6)**. Consistent with the fact that a starvation, or stringent-like response, is induced by CO_2 downshift we observe a lag phase in WT *B. ovis* growth upon re-introduction to a 5% CO_2 environment **(Figure 5A)**.

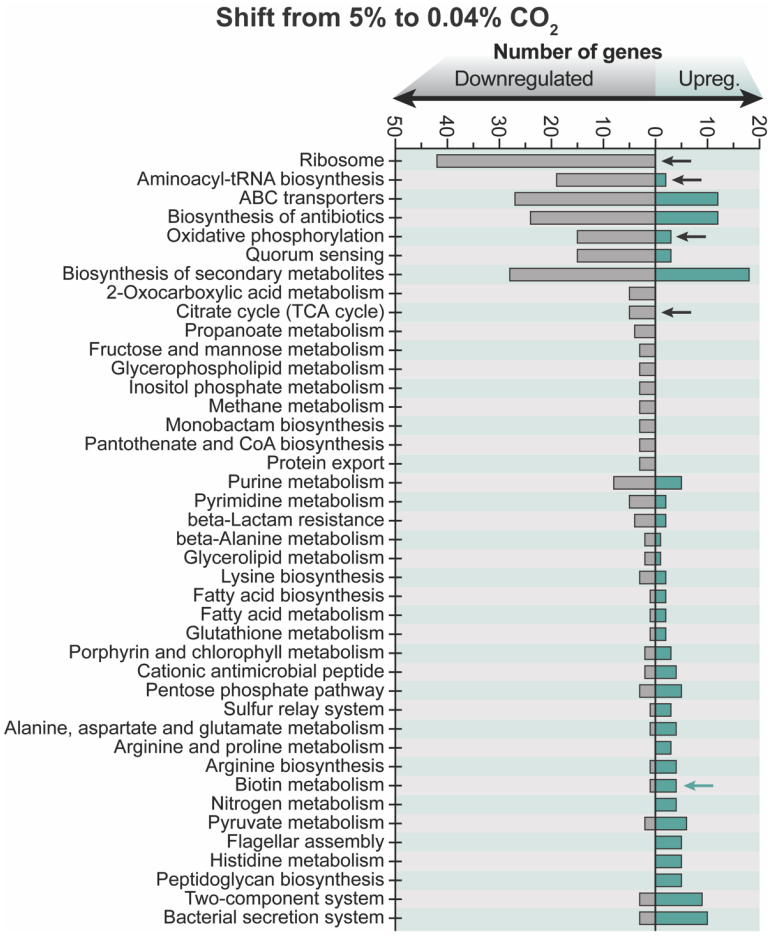


Figure 6. KEGG pathway assignment of genes regulated in wild-type *B. ovis* in response to CO_2 downshift

KEGG pathways are shown vertically on the left, with bars indicating the number of CO_2-regulated genes assigned to each pathway. Genes significantly upregulated in response to CO_2 downshift are on the right (**teal** bars), and genes significantly downregulated in response to CO_2 downshift are on the left (**gray** bars). Significance thresholds follow that outlined in **Figure 5B** legend. Pathways with less than 3 genes, or where the number of upregulated versus downregulated genes was not significantly different (see **Materials and Methods**), were removed from graph. **Black** arrows mark select KEGG categories associated with the stringent response, **teal arrow** marks the "Biotin metabolism" pathway.

Expression of genes implicated in response to stress, nutrient limitation, or stationary phase including polyphosphate kinases (*ppk1*, *ppk2*) (50-52), small DUF1127-family proteins (*BOV_RS04430*, *BOV_RS13510, BOV_RS13790, BOV_RS13950*) (53), and alternative sigma factors *rpoH* (54) and *ecfG/rpoE1* (55) is activated upon CO_2 downshift. Furthermore, Entner-Doudoroff/pentose phosphate pathway genes (*zwf, pgl, edd)* are transcriptionally activated, as are genes implicated in *Brucella* spp. virulence and infection including the *virB* type IV secretion gene cluster, the urease cluster, and several annotated flagellar genes **(Figs. 6 & 7 and Data Set 2)**. Although *B. ovis* is urease negative and possesses a degraded urease gene cluster (25), transcription of the *ure* genes (and pseudogenes) is activated upon CO_2

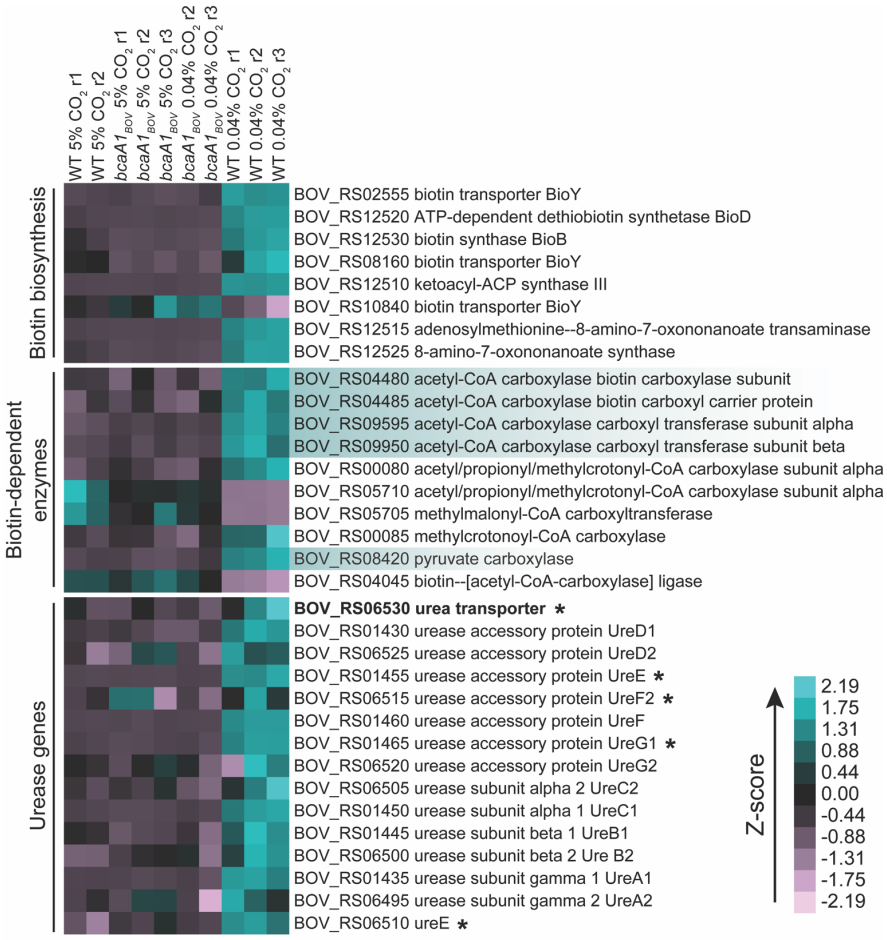


Figure 7. Heat map representation of a subset of genes regulated by CO_2 downshift in *B. ovis*
B. ovis gene accession numbers and RefSeq annotations are shown to the right of the heatmap. Genes were grouped into three categories: biotin biosynthesis, biotin-dependent enzymes and urease genes. Each column represents an individual RNA sample from either *B. ovis* ATCC 25840 (WT) or *B. ovis* $bcaA1_{BOV}$ grown in 5% or downshifted to 0.04% CO_2, measured by RNA-seq. Strain replicates are indicated r1, r2, and r3. The four genes that encompass the acetyl-CoA carboxylase complex and pyruvate carboxylase are highlighted in **teal**. *ure* pseudogenes are marked with an asterisk; the urea transporter, which is a pseudogene in *B. ovis*, is in **bold**. Z-scores express the number of standard deviations from the mean for each gene (see Materials and Methods).

downshift. Among the most highly activated gene sets in low CO_2 are biotin biosynthesis genes **(Figs. 6 & 7 and Data Set 2, 3)**. Biotin is a cofactor involved in carboxylation reactions that use CO_2 as a substrate. Concordant with the result that cells remodel transcription to enhance carboxylation reactions, we also observed that transcription of pyruvate carboxylase and several biotin-dependent acetyl-CoA carboxylases and carboxyl transferases is significantly enhanced **(Fig. 7 and Data Set 2)**.

Pseudogene sequences, including $bcaA_{BOV}$, are highly conserved across all *B. ovis* isolates

All 17 sequenced *B. ovis* isolates (**Data Set 1**) carry the same single nucleotide insertion in $bcaA_{BOV}$, yielding a frameshifted gene that apparently encodes a non-functional carbonic anhydrase protein. From these data, we have concluded that $bcaA_{BOV}$ is a pseudogene. Since pseudogenes are predicted to be under neutral selection (56-58), we might expect that $bcaA_{BOV}$ should accumulate mutations at a higher frequency than functional genes. Actually, there is evidence that pseudogenes are rapidly deleted from bacterial genomes due to either toxic side effects of expression, or the energetic burden of maintaining the gene (59). However, we observe no other mutations within $bcaA_{BOV}$ across all sequenced *B. ovis* strains, which have been isolated from regions across the globe over the past 60 years. This suggests either that the nucleotide insertion mutation at the 3' end of $bcaA_{BOV}$ occurred very recently in the evolutionary history of *B. ovis*, or that selective pressures maintain this pseudogene, raising the possibility that this gene sequence may have an additional function. It is known that $bcaA_{BOV}$ does not contribute to *B. ovis* infection in a mouse model of disease (27), but we have shown that this gene is transcribed **(GEO accession GSE130678 and Fig. 5E)** and thus could contribute other functions. To assess if the conservation of $bcaA_{BOV}$ allele in *B. ovis* differs from other pseudogenes, we defined the frequency of mutations and sequence divergence across a larger group of both annotated pseudogenes and functional genes within in *B. ovis* and compared to *B. abortus* strain 2308. Our aim was to use this gene set to empirically define pseudogene divergence in *B. ovis*, and to use this set as a point of comparison for $bcaA_{BOV}$.

We quantified polymorphisms in three distinct sets of genes; 1) urease genes, 2) the VirB type IV secretion system (T4SS), and 3) *B. ovis* pseudogenes described by Tsolis *et al.*(25) (**Table S2**). *B. ovis* is urease negative (23, 60) and the gene clusters shows features of degradation (25) with intact genes and pseudogenes that are presumably both under neutral selection as the pathway is non-functional. *B. abortus,* on the other hand, is urease positive and has an intact gene cluster (61). The T4SS provided us with

a set of genes that are functional and intact in both species (62-64). The additional *B. ovis* pseudogenes included in our analysis comprise multiple erythritol catabolic genes and cytochrome oxidase genes, which each encode full-length counterparts in *B. abortus* (**Table S2**). When assessing polymorphisms, each gap (insertion or deletion) was considered as a single event, regardless of its length.

We observed strikingly low divergence within *B. ovis* isolates. In other words, the sequences of all *B. ovis* pseudogenes in our analyzed set are highly conserved across a geographically diverse group of strains isolated between 1959 and 2011. Only 4 pseudogenes — *ureT* (BOV_RS06530), *norB* (BOV_RS11505), *ccoO* (BOV_RS01915) and *pckA* (BOV_RS09880) — exhibit polymorphisms between *B. ovis* strains, and each is polymorphic at only one site **(Table S2)**. Similarly, sequence differences were rare when comparing the *ure* genes or the T4SS genes within the 17 *B. ovis* strains. In short, the level of polymorphism we observe between the sequenced *B. ovis* isolates is not significantly different between a functional group of genes (*virB*) and a non-functional gene cluster containing pseudogenes (*ure*).

We then sought to compare these genes to *B. abortus* ATCC 2308. Overall, annotated pseudogenes have the highest average number polymorphisms per base pair between *B. ovis* strains and *B. abortus*. As expected, genes in the functional T4SS (*virB*) locus have the lowest average number of differences per base pair and a bias towards synonymous mutations (**Fig. 8**). *B. ovis* does not produce functional urease and the *ure* genes, accordingly, have a higher number of mutations compared to *B. abortus*. We do observe that larger or out-of-frame deletions are more prevalent in the *ure* genes than the *virB* cluster (**Table S2**) consistent with a model in which the *ure* genes are under relaxed selective pressure. Though pseudogenes are, on average, more highly diverged between *B. ovis* and *B. abortus* (**Fig. 8**), our analysis shows that many other pseudogene sequences besides $bcaA_{BOV}$ are also entirely conserved in *B. ovis* (**Table S2**). Thus, we cannot conclude that the conservation of $bcaA_{BOV}$ across *B. ovis* isolates is due to purifying selection as a result of $BcaA_{BOV}$ performing some other function in the cell. Rather, a more likely explanation for our results is that the loss of function mutation resulting in a *B. ovis* $bcaA_{BOV}$ pseudogene occurred very recently in the evolutionary history of an ancestral *B. ovis* strain that now commonly infects sheep across the globe.

Discussion

Brucella cultivation and the resurrection of a carbonic anhydrase pseudogene

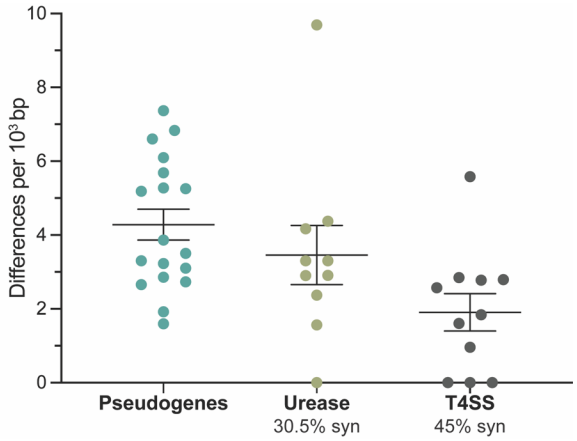


Figure 8. Nucleotide differences between sets of functional genes and pseudogenes in *B. ovis* ATCC 25840 and their *B. abortus* ATCC 2308 homologs normalized by gene length
Each dot represents a gene. Genes are cataloged in **Table S2**. Bars represent the mean and standard error of the mean for each group. Percentage of synonymous mutations in urease and T4SS genes are indicated.

Brucella spp. are intracellular mammalian pathogens that are highly related (65) but have clear host preferences (2-4) and, in some cases, distinguishing physiological features (26). During the course of transmission and infection, *Brucella* spp. must endure numerous shifts in the chemical state of the environment (3, 66), including changes in steady-state levels of CO_2. In the aqueous environment of the cell, CO_2 can spontaneously hydrate to form bicarbonate, an important substrate for microbial growth (67). The importance of bicarbonate to life is evidenced by the fact that at least 7 different families of carbonic anhydrases (28, 40, 68) have evolved across all kingdoms of life. These enzymes function to enhance the rate of CO_2 hydration to produce bicarbonate. *Brucella* spp. encode at least three carbonic anhydrases, but our data and recently published data (27) provide evidence that a β-class carbonic anhydrase encoded by *bcaA* is specifically required for axenic growth under the low partial CO_2 pressure (pCO_2) of a standard atmosphere in *B. ovis* and *B. abortus*.

While almost all *Brucella* spp. can be cultivated without adding CO_2, the majority of *B. abortus* strains and all wild *B. ovis* strains have a strict requirement for CO_2 supplementation for axenic cultivation. This requirement confounded early efforts to cultivate *B. abortus* from cattle (8, 10), and the knowledge that *B. abortus* strains require added CO_2 enabled early efforts to cultivate *B. ovis* from sheep tissue (23, 69). Our analysis of over 700 sequenced *Brucella* isolates of multiple species provides evidence that single nucleotide insertions in *bcaA*

result in frameshifts that lead to non-functional BcaA protein in *B. abortus* and *B. ovis*. Both of these species carry distinct insertion mutations in *bcaA* (**Figs. 1 & 4 and Data Set 1**) that underlie the strict CO_2 requirement for growth outside of an animal host. Functional, reverted alleles of these frameshifted *B. abortus* and *B. ovis bcaA* pseudogenes are easily selected by incubating cells in broth under standard atmospheric conditions for several days. This is consistent with historical reports of loss of the CO_2 requirement after repeated passages of *B. abortus* and *B. ovis* in air (11, 23), and is conceptually in line with a recent report in *E. coli* showing functional resurrection of a pseudogene involved in iron acquisition when cells are cultivated under selective conditions (70).

The fact that *B. ovis* and *B. abortus* lineages have independent insertion/frameshift mutations in *bcaA* that have apparently fixed in populations of each species suggests there is an adaptive advantage to loss of *bcaA* function. Our data show that *B. ovis* mutants harboring the functional $bcaA1_{BOV}$ allele grow significantly faster than wild-type at high partial CO_2 pressure (**Fig. 1E and 1G**), which is what *Brucella* spp. encounter in mammalian tissues. Given that slow growth can be advantageous for bacterial pathogens to maintain long-term infections (71, 72), it is possible that loss of *bcaA* function enhances *Brucella* fitness in the context of chronically-infected tissue in natural hosts. Notably, *B. ovis* is almost entirely restricted to male reproductive tissue and is reported to be transmitted only through venereal or direct contact (23, 73, 74). The rate of oxygen uptake and the partial pressure of CO_2 are higher in testis than in other tissues (75, 76), thus it is conceivable that *B. ovis* has evolved a lifestyle in which carbonic anhydrases are no longer required. Reduced growth rate as a consequence of *bcaA* frameshift may provide *B. ovis* with a fitness advantage during chronic infection in the high pCO_2 environment of the testis. *B. abortus*, on the other hand, can be transmitted sexually, through inhalation, or by ingestion (3, 77). Although there is no known *Brucella* niche outside a mammalian host, it is well established that bacteria on aborted tissues is one of the main routes of transmission for *B. abortus* (78). It would seem that a functional BcaA could be advantageous for CO_2 assimilation when *B. abortus* is residing on aborted tissue outside its animal host, though the extraordinarily high titer of respiring *B. abortus* on aborted fetal tissue (79) may lead to a high enough local pCO_2 that BcaA is unnecessary for growth in that context. Alternatively, growth may be simply impeded during this relatively narrow window outside the host. Indeed, growth impairment by CO_2 limitation outside the host could serve as an important signal for the bacterium. A minority of sequenced *B. abortus* isolates encode a full-length (and presumably functional) BcaA, and we predict that some of these isolates acquired gain-of-function mutations described here, during cultivation in the laboratory before they were sequenced (11).

CO_2 limitation elicits a starvation and virulence gene expression response

As outlined in the introduction, $^{14}CO_2$ was recovered primarily in pyrimidines and glycine in *B. abortus* (15, 80). Our efforts to bypass the *B. ovis* CO_2 requirement by supplementing media with amino acids, nucleotides, and bicarbonate were not successful, just as earlier efforts to bypass the *B. abortus* CO_2 requirement by adding bicarbonate or other lysates to the medium were unsuccessful (12). Thus, it seems likely that *Brucella* spp. lack a functional bicarbonate transporter. Our transcriptomic data provide clear evidence that shifting *B. ovis* to a standard atmosphere induces a starvation response that has features of the stringent response (46, 49, 81). This response is evidenced by the downregulation of transcription and translation related genes as well as genes involved in respiration and central metabolism (**Fig. 6**). Among the most significantly activated set of genes upon CO_2 downshift is the *virB* type IV secretion system, suggesting that cells initiate a gene expression program that could influence virulence when they encounter a CO_2 limited environment. Growth and transcription in a *B. ovis* strain with a functionally-restored *bcaA* gene (i.e. $bcaA1_{bov}$) was completely insensitive to a shift in pCO_2. We propose that loss-of-function frameshift mutations in *bcaA* yield strains that can more readily sense changes in pCO_2 in the host. Beyond slowing the growth of *B. ovis* under the high partial CO_2 pressure of the host environment, it is possible that loss of *bcaA* function also endows cells with an ability to more acutely detect changes in environmental CO_2 levels, for example when a *Brucella* cell is ejected from its animal host. Whether this change in the CO_2 detection threshold is advantageous in natural infection and transmission contexts remains to be determined.

Acknowledgments

We would like to thank Julien Herrou for invaluable experimental help and fruitful scientific discussions. We thank David Hershey for aid in Tn-Himar library construction and helpful scientific debates. We also thank the Crosson lab members for the scientific support received. This work performed with support from the NIH Genetics and Regulation Training Grant (L.M.V.), NIH R01AI107159 (S.C.), and the Gallo Global Health Fellowship Program (L.M.V.).

References

1. Seleem MN, Boyle SM, Sriranganathan N. 2010. Brucellosis: a re-emerging zoonosis. Vet Microbiol 140:392-8.
2. Godfroid J, Garin-Bastuji B, Saegerman C, Blasco JM. 2013. Brucellosis in terrestrial wildlife. Rev Sci Tech 32:27-42.
3. Moreno E. 2014. Retrospective and prospective perspectives on zoonotic brucellosis. Front Microbiol 5:213.
4. Whatmore AM. 2009. Current understanding of the genetic diversity of Brucella, an expanding genus of zoonotic pathogens. Infect Genet Evol 9:1168-84.
5. Alton GG, Forsyth JRL. 1996. Brucella. *In* Baron S (ed), Medical Microbiology, Galveston (TX).
6. Franco MP, Mulder M, Gilman RH, Smits HL. 2007. Human brucellosis. Lancet Infect Dis 7:775-86.
7. Al Dahouk S, Hofer E, Tomaso H, Vergnaud G, Le Fleche P, Cloeckaert A, Koylass MS, Whatmore AM, Nockler K, Scholz HC. 2012. Intraspecies biodiversity of the genetically homologous species Brucella microti. Appl Environ Microbiol 78:1534-43.
8. Bang B. 1897. Die Aetiologie des seuchenhaften („infectiösen“) Verwerfens. Zeitschrift für tiermedizin 1:241-278.
9. Henry B, Traum J, CM H. 1932. Methods for the isolation of *Brucella abortus*. Hilgardia 6:355-379.
10. Huddleson IF. 1921. The importance of an increased carbon dioxide tension in growing *Bact. abortus* (Bang). Cornell Vet 11:210-215.
11. Smith T. 1924. Some Cultural Characters of *Bacillus abortus* (Bang) with Special Reference to CO2 Requirements. J Exp Med 40:219-32.
12. Wilson GS. 1931. The gaseous requirements of *Br. abortus* (bovine type). Br J Exp Pathol 12:88-92.
13. Marr AG, Wilson JB. 1951. Fixation of C14O2 in amino acids by *Brucella abortus*. Arch Biochem Biophys 34:442-448.
14. Newton JW, Marr AG, Wilson JB. 1954. Fixation of C14O2 into nucleic acid constituents by *Brucella abortus*. J Bacteriol 67:233-6.
15. Tepper BS, Wilson JB. 1958. Fixation and distribution of C14O2 in *Brucella abortus*. J Bacteriol 76:24-8.
16. Gerhardt P, Wilson JB. 1950. Attempts to replace the added carbon dioxide required by some strains of *Brucella abortus*. J Bacteriol 59:311-2.
17. Bruce D. 1887. Note on the discovery of a microorganism in Malta fever. Practitioner 39:161-170.
18. Spink WW. 1956. The Evolution of the Concept that Brucellosis is a Disease of Animals and Man, p 3-28, The Nature of Brucellosis. University of Minnesota, Minneapolis, MN.
19. Evans AC. 1918. Further studies on bacterium abortus and related bacteria. J Infect Dis 22:580-593.
20. Evans AC. 1923. The nomenclature of the melitensis-abortus group of bacterial organisms. Public Health Reports 38:1943-1948.
21. Hardy AV, Jordan CF, Borts IH, Hardy GC. 1930. Undulant Fever: With Special Reference to a Study of *Brucella* Infection in Iowa. Public Health Reports 45:2433-2474.
22. Fitch CP. 1922. The cultivation of *Bacterium abortus* Bang. J Infect Dis 31:233-236.
23. Buddle MB. 1956. Studies on *Brucella ovis* (n.sp.), a cause of genital disease of sheep in New Zealand and Australia. J Hyg (Lond) 54:351-64.
24. Marr AG, Wilson JB. 1950. The carbon dioxide requirements of *Brucella abortus*. Third Inter-American Congress on Brucellosis 1:151-155.
25. Tsolis RM, Seshadri R, Santos RL, Sangari FJ, Lobo JM, de Jong MF, Ren Q, Myers G, Brinkac LM, Nelson WC, Deboy RT, Angiuoli S, Khouri H, Dimitrov G, Robinson JR, Mulligan S, Walker RL, Elzer PE, Hassan KA, Paulsen IT. 2009. Genome degradation in *Brucella ovis* corresponds with narrowing of its host range and tissue tropism. PLoS One 4:e5519.
26. Alton GG, Jones LM, Pietz DE. 1975. Laboratory techniques in brucellosis. Monogr Ser World Health Organ:1-163.
27. Perez-Etayo L, de Miguel MJ, Conde-Alvarez R, Munoz PM, Khames M, Iriarte M, Moriyon I, Zuniga-Ripa A. 2018. The CO2-dependence of Brucella ovis and Brucella abortus biovars is caused by defective carbonic anhydrases. Vet Res 49:85.
28. Supuran CT. 2016. Structure and function of carbonic anhydrases. Biochem J 473:2023-32.
29. Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol Biol 1151:165-88.
30. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. Nucleic Acids Res 45:D535-D542.

31. Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27-30.
32. Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, Hoover CA, Blow MJ, Bristow J, Butland G, Arkin AP, Deutschbauer A. 2015. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. MBio 6:e00306-15.
33. Herrou J, Willett JW, Fiebig A, Czyz DM, Cheng JX, Ultee E, Briegel A, Bigelow L, Babnigg G, Kim Y, Crosson S. 2019. Brucella periplasmic protein EipB is a molecular determinant of cell envelope integrity and virulence. J Bacteriol doi:10.1128/JB.00134-19.
34. Herrou J, Willett JW, Fiebig A, Varesio LM, Czyz DM, Cheng JX, Ultee E, Briegel A, Bigelow L, Babnigg G, Kim Y, Crosson S. 2019. Periplasmic protein EipA determines envelope stress resistance and virulence in Brucella abortus. Mol Microbiol 111:637-661.
35. de Hoon MJ, Imoto S, Nolan J, Miyano S. 2004. Open source clustering software. Bioinformatics 20:1453-4.
36. Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-8.
37. Demerec M, Adelberg EA, Clark AJ, Hartman PE. 1966. A proposal for a uniform nomenclature in bacterial genetics. Genetics 54:61-76.
38. Smith KS, Ferry JG. 2000. Prokaryotic carbonic anhydrases. FEMS Microbiol Rev 24:335-66.
39. DiMario RJ, Clayton H, Mukherjee A, Ludwig M, Moroney JV. 2017. Plant Carbonic Anhydrases: Structures, Locations, Evolution, and Physiological Roles. Mol Plant 10:30-46.
40. Sly WS, Hu PY. 1995. Human carbonic anhydrases and carbonic anhydrase deficiencies. Annu Rev Biochem 64:375-401.
41. Supuran CT. 2016. Legionella pneumophila Carbonic Anhydrases: Underexplored Antibacterial Drug Targets. Pathogens 5.
42. Joseph P, Ouahrani-Bettache S, Montero JL, Nishimori I, Minakuchi T, Vullo D, Scozzafava A, Winum JY, Kohler S, Supuran CT. 2011. A new beta-carbonic anhydrase from Brucella suis, its cloning, characterization, and inhibition with sulfonamides and sulfamates, leading to impaired pathogen growth. Bioorg Med Chem 19:1172-8.
43. Joseph P, Turtaut F, Ouahrani-Bettache S, Montero JL, Nishimori I, Minakuchi T, Vullo D, Scozzafava A, Kohler S, Winum JY, Supuran CT. 2010. Cloning, characterization, and inhibition studies of a beta-carbonic anhydrase from *Brucella suis*. J Med Chem 53:2277-85.
44. Kohler S, Ouahrani-Bettache S, Winum JY. 2017. Brucella suis carbonic anhydrases and their inhibitors: Towards alternative antibiotics? J Enzyme Inhib Med Chem 32:683-687.
45. Merlin C, Masters M, McAteer S, Coulson A. 2003. Why is carbonic anhydrase essential to Escherichia coli? J Bacteriol 185:6415-24.
46. Boutte CC, Crosson S. 2011. The complex logic of stringent response regulation in Caulobacter crescentus: starvation signalling in an oligotrophic environment. Mol Microbiol 80:695-714.
47. Durfee T, Hansen AM, Zhi H, Blattner FR, Jin DJ. 2008. Transcription profiling of the stringent response in Escherichia coli. J Bacteriol 190:1084-96.
48. Paul BJ, Ross W, Gaal T, Gourse RL. 2004. rRNA transcription in Escherichia coli. Annu Rev Genet 38:749-70.
49. Traxler MF, Summers SM, Nguyen HT, Zacharia VM, Hightower GA, Smith JT, Conway T. 2008. The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. Mol Microbiol 68:1128-48.
50. Gray MJ, Jakob U. 2015. Oxidative stress protection by polyphosphate--new roles for an old player. Curr Opin Microbiol 24:1-6.
51. Henry JT, Crosson S. 2013. Chromosome replication and segregation govern the biogenesis and inheritance of inorganic polyphosphate granules. Mol Biol Cell 24:3177-86.
52. Rao NN, Gomez-Garcia MR, Kornberg A. 2009. Inorganic polyphosphate: essential for growth and survival. Annu Rev Biochem 78:605-47.
53. Budnick JA, Sheehan LM, Kang L, Michalak P, Caswell CC. 2018. Characterization of Three Small Proteins in Brucella abortus Linked to Fucose Utilization. J Bacteriol 200.
54. Delory M, Hallez R, Letesson JJ, De Bolle X. 2006. An RpoH-like heat shock sigma factor is involved in stress response and virulence in Brucella melitensis 16M. J Bacteriol 188:7707-10.
55. Kim HS, Caswell CC, Foreman R, Roop RM, 2nd, Crosson S. 2013. The Brucella abortus general stress response system regulates chronic mammalian infection and is controlled by phosphorylation and proteolysis. J Biol Chem 288:13906-16.
56. Li WH, Gojobori T, Nei M. 1981. Pseudogenes as a paradigm of neutral evolution. Nature 292:237-9.
57. Li WH, Wu CI, Luo CC. 1984. Nonrandomness of point mutation as reflected in nucleotide

substitutions in pseudogenes and its evolutionary implications. J Mol Evol 21:58-71.
58. Petrov DA, Hartl DL. 2000. Pseudogene evolution and natural selection for a compact genome. J Hered 91:221-7.
59. Kuo CH, Ochman H. 2010. The extinction dynamics of bacterial pseudogenes. PLoS Genet 6:10.1371/journal.pgen.1001050.
60. Corbel MJ, Hendry DM. 1985. Urease activity of Brucella species. Res Vet Sci 38:252-3.
61. Sangari FJ, Seoane A, Rodriguez MC, Aguero J, Garcia Lobo JM. 2007. Characterization of the urease operon of Brucella abortus and assessment of its role in virulence of the bacterium. Infect Immun 75:774-80.
62. Hong PC, Tsolis RM, Ficht TA. 2000. Identification of genes required for chronic persistence of Brucella abortus in mice. Infect Immun 68:4102-7.
63. Macedo AA, Silva AP, Mol JP, Costa LF, Garcia LN, Araujo MS, Martins Filho OA, Paixao TA, Santos RL. 2015. The abcEDCBA-Encoded ABC Transporter and the virB Operon-Encoded Type IV Secretion System of Brucella ovis Are Critical for Intracellular Trafficking and Survival in Ovine Monocyte-Derived Macrophages. PLoS One 10:e0138131.
64. Sieira R, Comerci DJ, Sanchez DO, Ugalde RA. 2000. A homologue of an operon required for DNA transfer in Agrobacterium is required in Brucella abortus for virulence and intracellular multiplication. J Bacteriol 182:4849-55.
65. Wattam AR, Foster JT, Mane SP, Beckstrom-Sternberg SM, Beckstrom-Sternberg JM, Dickerman AW, Keim P, Pearson T, Shukla M, Ward DV, Williams KP, Sobral BW, Tsolis RM, Whatmore AM, O'Callaghan D. 2014. Comparative phylogenomics and evolution of the Brucellae reveal a path to virulence. J Bacteriol 196:920-30.
66. Celli J. 2015. The changing nature of the Brucella-containing vacuole. Cell Microbiol 17:951-8.
67. Blombach B, Takors R. 2015. CO2 - Intrinsic Product, Essential Substrate, and Regulatory Trigger of Microbial and Mammalian Production Processes. Front Bioeng Biotechnol 3:108.
68. Supuran CT, Capasso C. 2017. An Overview of the Bacterial Carbonic Anhydrases. Metabolites 7.
69. Buddle M, Boyes B. 1953. A *Brucella* Mutant Causing Genital Disease Of Sheep In New Zealand. Aust Vet J 29:145-153.
70. Anand A, Olson CA, Yang L, Sastry AV, Catoiu E, Choudhary KS, Phaneuf PV, Sandberg TE, Xu S, Hefner Y, Szubin R, Feist AM, Palsson BO. 2019. Pseudogene repair driven by selection pressure applied in experimental evolution. Nat Microbiol 4:386-389.
71. Gengenbacher M, Kaufmann SH. 2012. Mycobacterium tuberculosis: success through dormancy. FEMS Microbiol Rev 36:514-32.
72. Neubauer C, Kasi AS, Grahl N, Sessions AL, Kopf SH, Kato R, Hogan DA, Newman DK. 2018. Refining the Application of Microbial Lipids as Tracers of Staphylococcus aureus Growth Rates in Cystic Fibrosis Sputum. J Bacteriol 200:10.1128/JB.00365-18.
73. Picard-Hagen N, Berthelot X, Champion JL, Eon L, Lyazrhi F, Marois M, Peglion M, Schuster A, Trouche C, Garin-Bastuji B. 2015. Contagious epididymitis due to Brucella ovis: relationship between sexual function, serology and bacterial shedding in semen. BMC Vet Res 11:125.
74. Ridler AL, West DM. 2011. Control of Brucella ovis infection in sheep. Vet Clin North Am Food Anim Pract 27:61-66.
75. Free MJ, Schluntz GA, Jaffe RA. 1976. Respiratory gas tensions in tissues and fluids of the male rat reproductive tract. Biol Reprod 14:481-8.
76. Wenger RH, Katschinski DM. 2005. The hypoxic testis and post-meiotic expression of PAS domain proteins. Semin Cell Dev Biol 16:547-53.
77. Olsen SC, Palmer MV. 2014. Advancement of knowledge of Brucella over the past 50 years. Vet Pathol 51:1076-89.
78. Atluri VL, Xavier MN, de Jong MF, den Hartigh AB, Tsolis RM. 2011. Interactions of the human pathogenic *Brucella* species with their hosts. Annu Rev Microbiol 65:523-41.
79. Anderson TD, Cheville NF, Meador VP. 1986. Pathogenesis of placentitis in the goat inoculated with Brucella abortus. II. Ultrastructural studies. Vet Pathol 23:227-39.
80. Tepper BS, Wilson JB. 1958. A comparison of the amino acid composition and nucleic acid content of strains of Brucella abortus. J Infect Dis 103:19-24.
81. Boutte CC, Crosson S. 2013. Bacterial lifestyle shapes stringent response activation. Trends Microbiol 21:174-80.