THE UNIVERSITY OF CHICAGO

THE RESPONSE OF *BRUCELLA OVIS* TO CHANGING AND CHALLENGING ENVIRONMENTS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

COMMITTEE ON MICROBIOLOGY

BY

LYDIA MIMI VARESIO

CHICAGO, ILLINOIS

JUNE 2021

To my dad, who taught me to love science and was always there for me, supporting me in my endeavors and choices, I miss you so much; and to Luca Daveggio, for everything you meant to me, for all that I owe you, for the amazing person you were. You will always be in my heart.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Bacteria inhabit many environments, from hot dry desert regions to high pressure niches on the ocean floors, from high salt areas to cold latitudes, from inside plant root nodules to inside human cells. Studying the kinds of habitats different bacteria encounter and how they react to the challenges these environments offer is important to understand bacteria in their natural state, what kind of responses characterize them and how they can function when confronted with adversity.

Here I present my research on *Brucella ovis*, a facultative intracellular pathogen that has evolved to withstand stressful and harmful situations, from nutrient limitation, to oxidative stress, to drastic drops in pH. One ambient factor that perturbs *B. ovis* homeostasis is the atmospheric level of carbon dioxide, as *B. ovis* cannot be cultured in laboratory conditions without CO_2 supplementation. I examined the genetic underpinnings of the CO_2 dependence of *B. ovis* growth, identifying mutations in a carbonic anhydrase gene (*bcaA*) as responsible for this particular metabolic requirement. *B. ovis* harbors a unique, non-functional pseudogene allele of *bcaA* ($bcaA_{bov}$), and I found that some *B. abortus* lineages also harbor a *bcaA* pseudogene, thus rendering them dependent on CO_2 supplementation for growth. My data explain why *B. ovis* and select strains of *B. abortus* require elevated CO_2 levels for growth, which was first noted over a century ago. Transcription of one third of the genes in wild-type *Brucella ovis* change when cells are shifted from high to low CO_2 conditions; gene expression is unchanged upon CO_2 shift in a strain in which the pseudogene is restored to a functional *bcaA*. Thus, wild-type *B. ovis* has increased sensitivity to environmental levels of CO_2 because of a carbonic anhydrase pseudogene. This sensitivity could help *B. ovis* better detect when it is inside and outside the host.

I also present work in which I analyze *Brucella ovis* in the context of stationary phase, as a proxy to better understand the environment and related response that this pathogen encounters

within the host. I discovered that *cysE* -which encodes for a serine O-acetyltransferase, involved in the first step of de novo cysteine biosynthesis- is required for *B. ovis* fitness in stationary phase. Deletion of *cysE* increases sensitivity to hydrogen peroxide and attenuates *Brucella ovis* in tissue culture infection models. Thus, sulfur and cysteine metabolism is important for resistance to the hostile environment within the intracellular niche and presents an intriguing target for drug development against *Brucella* infections.

ACKNOWLEDGEMENTS

This dissertation is the celebration and culmination of more than five years of work. There are many people I would like to acknowledge as they have been pivotal to my success, provided much needed support, or have helped out in specific moments on particular issues where I needed it the most.

First of all, I would like to thank my PI and mentor, Sean Crosson. He has guided me for the past five years, teaching, coaching, steering and supporting me along the way. I have learned so much from this monumental person, and am exceedingly grateful that I got the chance to work, discover, and thrive in his lab. Being Sean's graduate student has made me a better scientist and gifted me the opportunity to freely explore my scientific interests, ask questions, discuss data and models, approach novel techniques and rediscover historical perspectives, and overall work in a positive, curious, and helpful environment he creates. Thank you mentoring me and all the time and patience you dedicated to my growth and learning.

I cannot think of the Crosson lab without Aretha. Aretha Fiebig is an amazing person and scientist. I'd like to thank her for the invaluable help and all the things she taught me throughout the years, for the long 'quick questions' that often derailed in topics more or less pertinent to the original query, for her understanding nature and endless patience, for her creative solutions and willingness to listen, for her openness in providing feedback and advice. She is one of my role models, both as a person and as a researcher, and has been fundamental to my betterment.

Julien Herrou was my scientific guardian angel for the longest time. He walked me through experiments and techniques, providing feedback and scientific discussion, mentoring me and dissuading me from falling into endless rabbit holes. He has been an amazing colleague and friend,

helping me stay focused and on track especially through the inevitable rough patches during my graduate career.

I would like to also acknowledge my former post doc mentor, Jonathan Willett, with whom I worked during my rotation in the Crosson lab and whose research in *Brucella* engrossed me, eventually leading to my joining the lab and picking up where he left off. His approachable manner and good nature eased me into the microbiology lab, which was an unfamiliar and novel environment for me at the time.

I would like to deeply thank David for the intense scientific discussions and teachings, for often challenging thoughts and results, therefore helping in the development of models, theories and data interpretation in more concrete, solid, and reasonable form, broadening my understanding and interpretation of data. And for being my honest friend, always happy to drink a beer and chat whenever I needed to. I would also like to thank Ben Stein, who moved with Sean, Aretha, and myself. Especially in East Lansing, Ben has helped me out both inside the lab and out, always lending an ear to let me vent and providing support and counsel, as well as investing time and energy in helping me out with scientific problems as well as day to day hassles. In East Lansing, I also found a friend and exceptional colleague in Patrick McLaughlin who has an amazing personality and whose strong will has fed well into mine, uplifting my mood and challenging my ideas in unique ways.

I would also like to generally acknowledge the Crosson lab, past and present members. It has morphed greatly during these five years, especially with the move, but there has always been inner lab support, fruitful discussions and advice during lab meetings as well as at the bench. From Chicago, I would like to specifically mention Daniel Eaton, for the Saturday lab laughs; Leila Ruiz-Reyes, for all the soccer games and comradery, and Maggie Zhang, for her shining words of

encouragement and pure niceness. I also want to thank the newer lab members that have helped ease the move and create a newly positive, rich, collaborative, and supportive environment in East Lansing, especially Hunter North, for her openness, positive energy, and for being an overall great person; Sergio Hernandez-Ortiz, for sharing music, games, and delicacies with me; and Tom Kim, for his serene presence and helpful manner.

I would like to deeply thank my Thesis Committee, for all the help they have provided in walking me through my graduate career, the scientific insights and advice, the encouragement and forward nudges: Glenn Randall, who believed in me from the start, and always has had time for me, with a smile; Maureen Coleman, who has always spontaneously offered me help and a place in her lab if needed; Lucia Rothman-Denes, whose honesty and affectionate manner have given me courage in the face of hard choices and solace in feeling supported and understood. I would also like to recognize my former Committee member, Howard Shuman, from whom I have learned so much, and whose wisdom and guidance were deeply appreciated.

Furthermore, I would like to acknowledge Matt Zurensky, Eric McLean, and Daniel Czyz, who helped me attempt a series of tissue culture experiments during my stay at the University of Chicago. I would also like to thank Gabriel Vargas, who has been at my side both scientifically, providing feedback, advice, and a constant willingness to chat about science, as well as for being an awesome, supporting, encouraging, and unwavering true friend.

I cannot imagine my time in Chicago without mentioning all the people that have been walking through this same career path with me, for its entirety or part of it, and have helped alleviate the harder moments by sharing games, sports, trips or just simple evenings and dinners, making this whole experience complete, joyous, and worthwhile: Shan Kasal, Chris Stamper, Andrew Miller, Audrey Williams, Alex Hoffman, Liana Hernandez, Zach Early, Ryan Duncombe,

Roy Morgan, Steven Erickson, Will Reidl, Matt Reyer, and Liz Ziegler. I would also like to mention a handful of friends that were not directly part my graduate cohort, that I met during orientation, that have unfailingly played pick up soccer with me, helping me vent my frustrations as well as providing cheer and merriment over nice cold beers: Marc Gillard, Paolo Andrich, Cem Randa, and George Adams. I would also mention Liz Lee and Alan Linton, for awesome broomball games and impressive comradery on those bitter cold nights.

There are a number of people that are not part of the University of Chicago that have nevertheless contributed to my success and happiness while pursuing my graduate degree, as well as helping me in the many years that lead up to the start of my PhD. The first person I want to thank, and that I can never thank enough, is my dad. There are no words to describe how much I owe him. He has taught me to love and appreciate nature, at all levels, as well as research, to ask questions, to be curious and search for answers. He has always prompted me to argue my points, discuss my ideas and opinions, to not shy away from learning new things and to hold my head high. He always wanted me to believe in myself, to trust in my instincts and knowledge, to be independent and strong. He has given me so much encouragement and support while also letting me walk on my own, make my mistakes, and grow from them.

I would also like to thank my group of lifelong 'genovese' friends that I left behind in Italy when I moved to the US. They have always been there for me. For long phone calls dealing with whatever problem in the universe we feel like discussing, to high end chats about politics, philosophy, sociology, art, and psychology to basic arguments about sex, drama, and gossip, to remembering trips, summers, and weekend outings, to talking about everyday problems, sad moments, frustrations, and angst, helping each other out how best we can and always, always ready to lend a hand, an ear, time. There are so many of you I would like to acknowledge, as you have

unwaveringly encouraged, supported and sustained over these years. Marc Bosi, Emiliano Montermini, Salvatore Canto, Giorgia Manavella, Davide Porzio, Martina Peloso, Francesco Parodi, Alessandro Gabbiani, Federica Camoirano, and Luca Villa, to name a few. Not from Genova, but still dear to my heart and just as important in these five long years, are those friends who are scattered throughout Piedmont, and who have given me amazing support and affection: Marco Ferrero, Camilla Fontana, Riccardo Pelle, Mauro Adriano, Sonia Palumbo, Danilo Monateri, and Jessica Dardano; as well as my colleagues and friends from my Master's lab, who contributed to my sanity and formation that eventually lead me to the Crosson lab, and have been willing participants to my rants and jests since: Teresa Poggio and Matteo Menotti. A special mention to Manuela Sigolo, the strongest person I know, genuine and compassionate, that has always reached out to me and cared for me at a distance. Thank you for everything.

I would like to finally acknowledge my family, first of all my mother, for teaching me strength and endurance, persistence and precision. For loving me and missing me, always taking care of me and for letting me know that I always have a home with her. I would also thank my sister, my blonde tall sister, for she is an incredible person, has forgiven my pigheadedness and has given me support and understanding, more than I deserved. I would like to thank the rest of my family, aunts, uncles and cousins, for the unconditional love and support you each have shown me all these years, the warm embrace and genuine smiles. A special thank you to my aunt Giovanna Eva, who is bitingly direct and loudly exuberant, and one of the closest confidants I have, who has aided me time and time again when I was in need, never flinching, never hesitating, always 100% there, listening, counseling, and uplifting. When I think of family, I cannot forget my new family, and I want to acknowledge the MacNabb family, especially Martha and Tom. They have embraced me as a daughter, taken care of me through the loss of my dad, knee surgery,

an unexpected move, and the many smaller hardships that these years have brought. They have celebrated my successes and have simply been two of the most generous, warm, affectionate people I ever had the fortune to meet. Thank you, so much.

Finally, I would like to thank Brendan MacNabb. Everything I thanked people up to now for, I also owe Brendan. He has been my anchor, my support in these wacky years, through the good and the bad. He has always stopped to listens without dismissing my concerns, has always tried to envision the best choices for me above his wishes, has given me all his support and held my hand through the many hardships. He has given me scientific advice. He has celebrated joyous days. He stood at my side during hard decisions. He took care of me when I needed it the most. He has understood my intense and impetuous nature. He is an incredible friend, a knowing colleague, a cunning teammate, an adventurous travel mate, and, obviously, an irreplaceable partner.

1. INTRODUCTION

All over our planet, bacteria inhabit an immensely broad collection of environments. They colonize a more diverse set of habitats than any other organism on earth. They have adapted –and are adapting– to their niches in a constant arms race with challenges set to them by that very niche: whether it be nutrient limitation, predation, dehydration, pH changes, osmotic shock, high pressure, hostile host attacks, radiation damage, or antibiotics (to name a few), they constantly have to retain the capability to fight or cope with these conditions. Bacteria rise to such challenges by sensing the offending environments and consequently adapting their physiology. These changes can be regulated and temporary (transcriptional, translational or post-translational) or may be permanent (i.e. genetic mutations). The latter may be then fixed within a population if it results in an overall fitness advantage for that population. There is a balance between plasticity and rigidity in the capacity of bacteria to survive and thrive in their environments, and different bacteria have different tools that enable them to survive their current environment and react to eventual changes. Thus, when cultivating bacteria in a laboratory, the conditions of growth and physiology examined often do not mirror the 'natural' living conditions encountered by that particular organism. Studying the nature of the different environments that bacteria face and how they respond to these conditions is paramount to understanding how bacteria thrive and interact, and allows us to delve deeper into what characteristics are important for them to flourish in their natural habitats.

In this dissertation, I describe how different environmental factors, specifically partial CO_2 pressures and sulfur-containing metabolites in the intracellular niche impact the fitness and possibly the evolution potential of the facultative intracellular bacterium, *Brucella ovis*. This

introductory chapter first presents key aspects of bacterial growth and some of the conserved responses that bacteria elicit upon different environmental perturbations. I next highlight characteristics of the Alphaproteobacteria, the class that *Brucella* belong to, before diving into details of *Brucella* physiology and infection biology. Since my thesis research involved investigation of host-derived stressors that impact *Brucella*, I will briefly cover a few details of the kinds of attacks *Brucella* face both within the host and specifically within the intracellular environment, and how this furtive microbe counteracts these challenging situations. As cysteine metabolism is one of the key features that enables *Brucella ovis* to survive host challenges, I will also detail aspects of cysteine and sulfur metabolism in my introduction. I will further describe *Brucella ovis* and highlight physiologic features that set it apart from other *Brucella* species, outlining why these characteristics are important and warrant further exploration. Finally, I will describe the importance of carbon dioxide in *Brucella* metabolism, the enzymes involved in CO_2 assimilation and their relationship to essential metabolic processes.

1.1. PREFACE

Some of the following introductory content has been adapted from two of my publications:

- **Varesio LM**, Willett JW, Fiebig A, Crosson S., *Journal of Bacteriology* (2019)
   DOI: 10.1128/JB.00509-19;
- **Varesio LM**, Fiebig A, Crosson S., *Infection and Immunity* (2021)
   DOI: 10.1128/IAI.00808-20.

1.2. BACTERIAL GROWTH

Bacteria possess sophisticated systems to sense their environment, and equally sophisticated methods to respond to it. They are rarely in ideal, nutrient rich situations like those we fabricate in laboratory conditions. Measuring bacterial growth is a good proxy to assess if the conditions they are in are well tailored to the organism in question.

One of the most obvious struggles for bacteria is to survive starvation. Nutrient scarcity is a common daily challenge that bacteria face. Nutrient limitation may be caused by natural environmental paucity or by sequestration from more fit competitors. When grown axenically in batch culture (i.e. where the nutrients are added at the start of the experiment), the depletion of available nutrients shapes the growth curve, which is divided into several phases. Initially, the cells are in lag phase, whose duration varies based on the conditions the bacteria were in prior to commencement of the growth experiment as well as the species/strain in question. The lag phase is followed by the log phase. During this second phase, bacteria are doubling exponentially. Physiologically, as they are actively growing and dividing, they favor energy-demanding reactions to synthesize nucleic acids, proteins, and other cell components. Once nutrients start to run out, or in response to other environmental changes that suppress growth (for instance a shift in pH), cells enter stationary phase (1). This phase is characterized by a gradual reduction of growth rate to zero and no net increase in cell numbers. Cells can persist in stationary phase for varying degrees of time in a more or less dynamic state, which depends on a multitude of factors, such as the starvation conditions (2). Eventually, cells pass from stationary phase to the fourth state, known as the death phase. Here, cells are rapidly dying, as can be assessed by colony forming unit (CFU) counts of viable cells, and up to 99% of cells in the population may perish. Whether this occurs stochastically or by programmed cell death (bacterial apoptosis) is not fully understood, but there is a small percent of the population that persists. These cells are now in what is known as the long-term

stationary phase, where a small population of viable bacteria can be detected. This is a dynamic state, where cells may die and consequently feed sister cells that can undergo some replication cycles or simply lay dormant without perishing. During this last phase, there is a particular signal phenotype associated with a series of morphological, physiological and transcriptional changes collectively known as the growth advantage in stationary phase (GASP) that allow cells to survive stress and persist in this final phase even for years (**Figure 6.1**) (3). The sensory pathways and responses needed for bacteria to survive in a persister state is poorly understood and currently an active area of study.

1.2.1. Stationary phase

Stationary phase is the period during growth in batch culture in which exponentially dividing cells slow growth and reach a plateau where there is no net increase in cell numbers. Stationary phase was initially studied in the model organism *Escherichia coli*, and entry into stationary phase has often been studied in the context of nutrient starvation (1, 2). Upon entry into stationary phase, cells undergo morphological changes in cell shape and size (4), increased resistance to heat shock and hydrogen peroxide (H_2O_2) (5), high tolerance to elevated osmolarity (6), release of nucleobases in the environment (7), decrease in protein synthesis, drastic metabolic shifts from biosynthetic metabolism to energy conservation, and overall global transcriptional, translational and post-translational modifications (2).

Studies in *E. coli* identified major central players of stationary phase induction. In 1990, Lange and Hengge-Aronis discovered an alternative sigma factor, named sigma S (σ^S, for stationary phase, or starvation) encoded by *rpoS* (8). Transcription (8) and stabilization (9) of RpoS is induced upon entry into stationary phase (i.e. reduction of growth rate), which in turn affects transcription of an immense regulon – more than 1000 genes in *E. coli* (10–12). RpoS regulation

occurs at multiple steps (13). *rpoS* is regulated at the transcriptional level by growth rate (and in some form by ppGpp, **see below**) and cyclic adenosine monophosphate (cAMP), an important second messenger (14). Another major player in regulating entry into stationary phase is the *hfq* gene, which encodes for host factor protein 1 (HF-1 or Hfq) (15). Hfq is an RNA-binding protein that was shown to affect RpoS levels by binding to the sigma factor in *E. coli* (16), indicating that Hfq is upstream of RpoS. It is now appreciated that Hfq is a global post-transcriptional regulator, affecting targets other than RpoS. Other molecules have been found to regulate transcription of *rpoS* in *E. coli*, such as H-NS (which inhibits Hfq and consequently blocks translation of *rpoS* mRNA) and OxyS (which also affects translation of *rpoS* by affecting the ribosome binding site) (17). RpoS is not the only regulatory molecule that regulates transcription during stationary phase. Depending on what commenced entry into this growth phase, other *rpoS*-independent proteins may come into play. For instance, other alternative sigma factors seem to play an important role, like *rpoH* (σ^{32} or σ^{H}) or *ecfG* (SigT) (**see below**) (1).

Stationary phase is not only induced by nutrient limitation. Accumulation of toxic metabolites, osmotic shock (18), various stresses (such as oxidative stress, **see below**) may also drive entry into stationary phase, although nutrient starvation is the most extensively studied. These different cell states are interconnected and partially overlap, based on the signals involved and on the bacterial species, allowing bacteria to react to different environmental cues and resist various stressful situations.

1.2.2. Stringent response

The stringent response is a conserved starvation response in bacteria. Different kinds of starvation, such as amino acid, fatty acid or, carbon depletion can trigger the stringent response, as well as osmotic stress (18–21), and they vary depending on the organism in question. The

immediate response is the rapid biosynthesis of the alarmone guanosine tetra- or penta-phosphate, (p)ppGpp (22). In *Escherichia coli*, where the stringent response was first studied, RelA or SpoT are responsible for synthesizing (p)ppGpp in response to different kinds of nutrient limitation (lack of amino acids in the case of RelA; fatty acid, iron, and carbon starvation in the case of SpoT). RelA binds RNA polymerase (RNAP) and, sensing the accumulation of uncharged transfer RNA (tRNA) molecules in the A site of RNAP as consequence of amino acid starvation (23), synthesizes (p)ppGpp. SpoT responds to different kinds of starvation, such as carbon, iron, and fatty acid limitation and also possesses a hydrolytic domain, so it is responsible not just for (p)ppGpp synthesis but also for its turnover (24). Bacteria have evolved varied mediators of the stringent response; in the Alphaproteobacterium *Brucella* (25, 26), for instance, a *relA*/*spoT*-homolog (*rsh*) replaces RelA and SpoT, and is the sole driver of (p)ppGpp synthesis and degradation.

(p)ppGpp was originally considered a master downregulator, as it negatively affects the half-life of most promoters thus repressing transcription, and in *E. coli*, increases levels of the alarmone to induce a decrease in stable RNA (tRNA ribosomal RNA (rRNA)) transcription as well as inhibition of DNA synthesis (27). It is now understood that (p)ppGpp is not only a negative regulator. Indeed, it also promotes transcription of *rpoS* (9) as well as of amino acid biosynthesis genes (28). Further studies have shown that in *E. coli*, more than 500 genes are differentially expressed in response to (p)ppGpp, the majority of which are upregulated (29). Of note, not all responses to the induction of the stringent response are the same in all bacteria. For instance, *Caulobacter crescentus*, another α-proteobacterium, does not react to amino acid starvation by synthesizing the alarmone (30), but activates the stringent response when glucose (31) and ammonium (32) are limmiting. Although the stringent response has been mostly characterized within the Proteobacteria phylum, recent studies have also uncovered the effects of the stringent

response in other groups of bacteria. For instance, the Gram-positive *Staphylococcus aureus* encodes for the RelA/SpoT homolog, RSH, as well as other two monofunctional synthetases, RelP and RelQ, which respond to cell-wall-targeting antimicrobials (33). In this organism, (p)ppGpp does not bind RNAP, but instead interacts with HprT/Gmk (enzymes involved in GTP synthesis) and inhibits their function, which in turn leads to a decrease of the levels of GTP and subsequent activation of CodY (a transcriptional repressor) and decrease in rRNA synthesis (34).

Thus, the stringent response is a highly conserved starvation response in bacteria which allows for efficient reaction to diverse environmental starvation cues. It is one prime example on how bacteria can modify their metabolism and behavior to react to external pressures that come from their habitat.

1.2.3. The general stress response

Stressful environments can also induce entry into stationary phase. Upon detection of a stressor, bacteria can activate what is known as the general stress response (GSR). This is a global transcriptional response which allows for increased survival in the presence of multiple stressors. The GSR is regulated by alternative sigma factors and can vary depending on the bacterial species.

In Gram-positive bacteria, SigB (σ^B) is the master regulator of the GSR, affecting the transcription of about 150 stress-related genes (35), and has been extensively studied in the model system *Bacillus subtilis*. The stress signals that control σ^B activity include both internal signals, such as the metabolic and energetic state of the cell (for instance, stationary phase entry signals (36)), and external signals such as low pH levels, and ethanol or heat shock (37, 38).

In Gram-negative bacteria, specifically in most of the Proteobacteria, σ^S is responsible for the modulation of the GSR. It has been extensively studied in the model organism *Escherichia coli*

(a δ-proteobacterium), where RpoS has been shown to respond to a wide array of diverse stressors, including low pH, heat shock, UV, osmotic stress and intracellular growth signals (11).

Not all Gram-negative bacteria encode for *rpoS*. Indeed, the GSR in the α-proteobacteria is instead tied to another alternative sigma factor: the extracytoplasmic function sigma factor EcfG (also named σ^T) (39). Since the members of the α-proteobacteria class live in diverse environments (**see below and Appendix A**), they activate the GSR in response to a wide array of stress signals, which include temperature shifts, desiccation, stationary phase, variation in pH levels, and UVs (40–42).

1.3. THE α-PROTEOBACTERIA CLASS

Proteobacteria comprise the largest bacterial phylum. This phylum is divided into nine classes, including the aforementioned α-proteobacteria and the γ-proteobacteria. The α-proteobacteria are one of the most diverse groups of bacteria and can be isolated from the most wide-ranging environments. For instance, the representative member of the Caulobacterales order, the asymmetrically dividing *Caulobacter crescentus*, was isolated from a pond in California (43); the aerobic anoxygenic phototroph (AAP, see **Appendix A**) *Erythrobacter litoralis* (of the Sphingomonodales order) was found in the cold salt waters off the island of Texel, in the Netherlands; the type species of the *Rhizobium* genus (of the order of the Rhizobiales), *Rhizobium leguminosarum*, instead grows in the roots of plants and in soil. Members of this class are also commonly isolated from the deep ocean floors (44) as well as from volcanic regions (45). Furthermore, there is a large diversity in genome architecture and size within this class and even within the same order: some genomes, like that of *Pelagibacter ubique* (from the *SAR11* clade) is one of the most compact genomes, with an average intergenic distance of 3 bp and a genome of

only 1.3 Mbp, making it the smallest known genome to date of this class (46); others, like *Mesorhizobium loti*, are much larger, about 7 Mbp in size (47), all the while sharing the same order – i.e. Rhizobiales – with *Bartonella quintiana*, whose genome size is only 1.7 Mbp (48).

The Alphaproteobacteria class also includes numerous unique pathogens. Among these are obligate intracellular pathogens like *Rickettsia*, which have a reduced genome (1.3 Mbp) due to permanent association with eukaryotic host cells, and the facultative intracellular pathogen *Brucella*, which possesses two chromosomes (total of 3.3 Mbp). In fact, bacteria that are in perpetual contact with eukaryotic cells (such as the obligate Alphaproteobacterial symbionts, *Wolbachia*) tend to have lost those genes that are no longer relevant to their survival, for instance biosynthesis genes for various metabolites. This is due to their nutritionally rich environment (i.e. the eukaryote host, cytoplasm) from which they can import vitamins and nutrients instead of having to expend energy in biosynthesis. On the other hand, those bacteria that live in nutrient-poor, changing environments are often more metabolically versatile, which enables them to adapt to more diverse nutritional conditions (49–51).

Thus, there is great variability within the alpha clade of proteobacteria, and they are uniquely useful model systems to study a variety of different morphological, genetic, environmental and evolutionary questions. The richness of the diverse environments they inhabit and the horde of distinctive challenges they face provides an outstanding reservoir of information to learn.

1.4. THE PATHOGEN *BRUCELLA*

Brucella are Gram-negative facultative intracellular pathogens, and members of the α-proteobacteria class. In 1887, David Bruce isolated the causative agent for a severe human disease

known as Malta fever, Mediterranean fever or Undulating Fever (now known as brucellosis), naming the pathogen *Bacillus melitensis* (now *Brucella melitensis*) (52). Subsequently, in 1897, Bernhard Bang discovered the etiological agent responsible for abortion in cattle, and named it *Bacillus Bang* (now *Brucella abortus*) (53). In fact, bacteria of the genus *Brucella* are the etiologic agents of brucellosis, which is among the most common zoonotic diseases worldwide (54, 55). It impacts human health (56) as well as the economy of affected regions as the most common practice to date to deal with an infected animal is to cull the entire herd (57). *Brucella* spp. can infect a range of wild and livestock animals (58), but have a relatively narrow host range and varying zoonotic potential (59, 60) and pathogenicity in humans. *B. melitensis*, for example, is primarily a sheep and goat pathogen and is often considered the most virulent species in humans (61). 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 (62). Considering that genetic identity across the genus is 94–98% at the coding level (60, 63), 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.

To date, there are 12 identified members of the *Brucella* family. They have been historically named according to the host they were originally discovered in and this usually reflects their preference in host species. The first three members identified were *Brucella melitensis*, *Brucella abortus* and *Brucella suis*, which preferentially infect ovine, bovines and swine, respectively. These bacteria all form smooth colonies on agar plates, which indicates they form a full lipopolysaccharide (LPS) layer (**see below**), are highly virulent, and have zoonotic potential.

There are only two recognized rough species of *Brucella* (i.e. lacking the O-chain of their LPS): *Brucella canis* (first described in 1966) (64) and *Brucella ovis* (discovered in 1953) (65), which were isolated from dogs and sheep, respectively. Other members of the *Brucella* clade include *Brucella neotomae* (found in desert wood rats, 1957) (66), *Brucella pinnipedialis* (1996) (67) and *Brucella ceti* (1994) (68, 69), both isolated from marine mammals, *Brucella microti* (found in voles, 2007) (70), *Brucella inopinata* (discovered in a breast implant, 2010) (71), *Brucella papionis* (in baboons, 2014) (72), and *Brucella vulpis* (isolated from a red fox, 2016) (73). While most *Brucella* were isolated from mammals, recent reports have identified novel *Brucella* spp. in amphibians (74, 75), thus increasing the range of possible hosts for this pathogen.

1.5. PATHOGENESIS

Brucella is a highly infective pathogen, its infectious dose varying based on the host and the route of entry, from tens to thousands (76) of bacteria. It can penetrate the human host via inhalation, ingestion, or direct contact with mucosa or damaged skin, is phagocytosed by professional (mostly dendritic cells (DCs) and macrophages) and non-professional phagocytes and then disseminates within the host (56, 77). *Brucella* infection usually progresses in three stages: an initial incubation stage, where there is no clinical evidence of infection and *Brucella* has passed the mucosal layer and entered the host; an acute phase, characterized by high numbers of *Brucella* disseminating through the host; a chronic phase, where the pathogen persists indefinitely, leading to organ damage and even death (78). In humans, the hallmark for brucellosis is the presence of an undulant fever, though this disease can present with a wide variety of symptoms (62).

Brucellosis is primarily an animal disease, and human to human transmission is a rare event (59, 79). In the animal host, disease progression follows a different course than in the human

patient (80). It can infect the host via the gastrointestinal tract, sexual contact or inhalation. From the local lymph nodes, it spreads mainly through the reticuloendothelial system, often initially targeting organs like the kidneys and spleen, though the exact dissemination pattern strongly depends of the host. It also has strong tropism for male genital organs, mammary glands, and placental trophoblasts, in the case of a pregnant host. Indeed, abortion is a hallmark of animal brucellosis and the aborted fetus is an important source of contagions for livestock, wild-life and humans, although animals can also initially present with an undulant fever.

Brucella is described as a stealthy pathogen. This is because it is particularly apt at hiding from the immune system, particularly the innate immune response (81). Compared to other enteric pathogens, whose aggression is accompanied by strong clinical effects such as diarrhea and high aggressive fevers, *Brucella* only elicits a mild innate immune response, and, in fact, the first clinical and characterizing symptoms of brucellosis in humans is an undulant fever often accompanied by abdominal pain. Indeed, strong activation of the innate immune system involves pro-inflammatory cytokines (TNF-α, IFN-γ, and IL-1β), and the presence of leukocytes (especially neutrophils) at the site of inflammation, as well as in the stool and blood, which are not a characteristic of brucellosis. This furtive pathogen effectively counteracts immune activation by many means. For instance, it can interfere with DC maturation through action of *Brucella*-TIP protein 1 (Btp1), thus dampening immune response and hindering the production of TNF-α and IL-12, and it does not bind the C3 component of the complement, which is another innate immune defense that attacks invading pathogens (78, 81–86).

Brucella also express an altered form of LPS, which is a pathogen-associated molecular pattern (PAMP) that is recognized by membrane-bound external toll-like receptor 4 (TLR4). LPS-from Gram-negative pathogens typically elicit a strong innate immune response, via activation of

the downstream pathways of TLR4. TLR4 triggers two main signaling pathways: the MyD88-dependent and the MyD88-indepentent pathways. The first leads to the activation of two transcription factors, NF-κB and AP-1, which have roles in the expression of pro-inflammatory cytokines. The MyD88-independent pathway activates the IRF3 transcription factor which induces production of type 1-interferons, such as IFN-α, which help regulate the activity of the immune system (87). Reduced TLR4 activation on the onset of *Brucella* infection means that there are only mild increases of pro-inflammatory cytokines like TNF-α and IL-1β, although some activation does occur at later stages of infection (82, 88, 89).

Furthermore, *Brucella* have evolved flagellin that lacks a domain that has been shown to elicit recognition by TLR5. TLR5 is another membrane-bound external TLR that specifically recognizes bacterial flagellin and activates NF-κB, AP-1, and IRF3 in response to the presence of this foreign protein (90, 91). Flagellin is the protein that constitutes the flagellar protofilament. *Brucella* are described as a non-motile pathogens (61) and there has been no evidence of the presence of chemotaxis genes or motility, with the exception of a *Brucella* species that was isolated from a Pac-man frog (*Ceratophrys ornata*) (74). This is surprising as most *Brucella* species encode the whole host of flagellar genes (92). Of note, there are reports describing the expression of a flagella-like structure within a sheath, which has been linked to virulence in *Brucella suis* (93). What the exact role of flagellin and flagellar encoding genes is in the context of brucellosis remains to be discerned.

1.6. THE INTRACELLULAR NICHE

Once *Brucella* is inside the host, it requires uptake by professional and non-professional phagocytes to establish infection (83). Only about 10% of *Brucella* survive phagocytosis (78, 94, 95). Nevertheless, this is an essential step in bacterial colonization.

Within the host cell, *Brucella* reside in what is known as the Brucella Containing Vacuole (BCV) which will mature from early (eBCV) endosome to late endosome and eventually fuse to the lysosome, leading to a drastic drop in pH (96). It appears that the endosomal compartment physiologically matures, unperturbed by the presence of *Brucella*. Indeed, it has been shown that it initially displays early endosome markers (like Rab-5 and EEA-1), then late endosome markers (such as CD63 and Rab7) and finally lysosomal markers, including LAMP1 and LAMP2, and bears the mark of a phagolysosome: pH levels of about 4-4.5. This happens within the first 8 hours post entry, with the lysosomal fusion occurring after the first 4 hours (97).

The lysosomal fusion is a fundamental step in *Brucella* host-cell infection. The drop in pH constitutes a signal for a set of virulence factors, most importantly the type IV secretion system (T4SS) operon, *virB*, whose expression peaks at 4 hours post-infection. T4SS is a secretion system which pumps effectors in the host cytosol (98). These effectors help promote further BCV maturation to the replicative BCV (rBCV), which is established between 8 – 12 hours post infection. At this stage, the eBCV loses endosomal markers and gains ER-related ones, such as calreticulin and calnexin. The T4SS effectors are still largely uncharacterized, but they seem to allow the eBCV to interact with the ER exit sites (ERES), with COPII (which coats vesicles trafficked from the rough ER to the Golgi) and Sar1 (a small GTPase that control COPII-mediated trafficking), among others, thus promoting the formation of the rBCV (99). The ER-derived vacuole is now an ideal niche for *Brucella* to replicate within the host cell, even though there are reports of replication initiation occurring before the rBCV has been completely established (97).

As the infection continues, the rBCVs fuse together in vast membranous structures that contain large numbers of bacteria. These vacuoles, known as the autophagosome BCV (aBCV), do not actually display autophagosome markers (like LC3) but rather late endosome and lysosome ones (such as LAMP1 and acidification). Autophagy is a cellular process that allows for internal degradation of cytoplasmic contents, whether from foreign invaders or from damaged intracellular proteins or organelles and does not appear to be completely activated by *Brucella*. Rather, a subset of autophagy-related proteins (such as Becklin1 and ULK1) are required for aBCV formation. *Brucella* can now egress from the cell (although this process is poorly understood) presumably by pirating the secretory functions of the autophagocytic pathway, and infect new cells. Of note, this process is non-lytic, and some *Brucella* maintain residence within the original host cell (100) while others egress and further disseminate within the host.

Within the host cell, *Brucella* is thus faced with a challenging and changing environment (101). First of all, it encounters nutrient limitation, which is known to induce entry into stationary phase (102), until the rBCV is established. Furthermore, *Brucella* can be subject to a direct attack from reactive oxygen species (ROS) and reactive nitrogen species (RNS) (103), often associated to the respiratory burst (104, 105) (**see below**), a conserved host cell response usually aimed at attacking invading pathogens (106). Additionally, as mentioned above, *Brucella* have to withstand drastic drops in pH levels once the lysosome fuses to the vacuole, and they are further subject to attacks by antimicrobial peptides as well as nitrosative stress. Some of these stressors are described in more detail below.

1.6.1. Nutrient limitation within the BCV

Once inside the macrophage and before the establishment of the rBCV, *Brucella* experience nutrient limitation. As described above, nutrient limitation is one of the factors that can

lead to entry into stationary phase which then becomes an interesting physiologic state to study how *Brucella* may behave within the intracellular environment. Indeed, the similarities of the two, stationary phase and the intracellular niche, have been nicely reviewed by Martin Roop and colleagues (102). For instance, *hfq* in *Brucella* has been shown to confer resistance in late stationary phase in *B. abortus* and to play a role in in vivo BALB/c mouse infection studies (107), and there are a number of genes in *Brucella* that are dependent on Hfq function, including *sodC* (which encodes for a superoxide dismutase, **see below**), that are important in *Brucella* in stationary phase (102). In 2006, the stringent response RelA/SpoT homolog *rsh* (**see above**), was shown to be important for *virB* expression and thus has an impact in virulence in *Brucella* (25), and in 2013 Hanna and colleagues studied the Rsh-dependent transcription profile in *Brucella suis* via microarray. Among others, they found that genes involved in stress adaptation (for instance *sodC*), protein folding, and chaperones were upregulated, whereas genes involved in amino acid and protein metabolism were largely downregulated. Expression of the transcriptional regulator *hutC*, which is a co-activator for *virB*, was also found to be upregulated in an Rsh-dependent manner (108) .

1.6.2. The respiratory burst and oxidative stress

Professional phagocytes have evolved to fight off invading pathogens. One strong defense mechanism they employ is known as the respiratory burst (109). When phagocytes sense pathogen invasion, they attack by secreting various species of oxidants. These oxidants include ROS (110), reactive nitrogen species (RNS), and reactive chlorine species (RCS) (103, 111, 112). They can also be produced endogenously (i.e. through aerobic respiration, in the presence of free metals and erroneously oxidized non-respiratory flavoproteins) (106).

Whether encountered in the environment or produced during growth, bacteria must avoid accumulation of high levels of these oxidants (113). One important gene involved in protecting against ROS is superoxide dismutase (SOD), which was first characterized in 1969 (114). SOD is a metalloenzyme (either Cu, Zn, Mn) that detoxifies the cells from the superoxide anion ($O_2^{\bullet-}$) producing H_2O_2 ($O_2^{\bullet-} + O_2 + 2\ H_2O \rightarrow H_2O_2 + O_2$). Hydrogen peroxide may react with metal ions such as iron(II) to produce a hydroxyl radical ($OH^{\bullet}$) (this is known as the Fenton reaction, $Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^{\bullet}$) (115). There are various enzymes that cells produce to decompose H_2O_2 into water and oxygen ($2\ H_2O_2 \rightarrow 2\ H_2O + O_2$), namely catalases (116), glutathione peroxidases (117) and peroxiredoxins (118). In the case of RNS, the best characterized defense mechanisms include flavohemoglobins (Hmp) and flavorubredoxins (NorV). These enzymes detoxify $NO^{\bullet}$ to nitrate ($2\ NO^{\bullet} + 2\ O_2 + NADPH \rightarrow 2\ NO_3^- + NAD(P)^+ + H^+$) in the case of Hmp (119), or to nitroxyl ($3\ NO^{\bullet} + NADPH \rightarrow NO^- + N_2O + NAD(P)^+ + H^+$), in the case of NorV (120).

Bacteria basally express scavenging systems as well as antioxidants to survive toxic oxidant accumulation. Among the various antioxidants, cysteine, methionine and glutathione are the most common. For instance, the thiol (-SH) group on cysteine's side chain is a strong nucleophile, which reacts readily with electrophilic species, producing sulfenic acids (-SOH) in the presence of two-electron oxidants (such as peroxides) or thiyl radicals ($RS^{\bullet}$) if the electrophile donates a single electron. If the thiyl radical reacts with a hydroxyl radical, the product is the highly reactive sulfenic acid (-SOH), which can either form disulfide bonds with a proximal cysteine or be irreversibly oxidized to sulfinic ($-SO_2H$) and sulfonic acid ($-SO_3H$) (121). Methionine residues can also be oxidized to methionine sulfoxides (Met-O) and rarely, Met-O is further and irreversibly oxidized to methionine sulfone (122).

Bacteria can also protect themselves from the oxidative stress, for instance, by storing cysteine in LMW thiols, thus protecting against metal-catalyzed autooxidation, and by sequestering metal ions (such as $Fe2^{+}$) in ferritins, bacterioferritins and Dps proteins, so as to reduce the chance of Fenton reactions taking place. Glutathione (GSH) is a low molecular weight (LMW) thiol that contains a reactive sulfhydryl group. It acts as a redox buffer during oxidative stress and helps maintain a reduced cytoplasm. It is found in plants, animals, fungi, in most Gram-negative bacteria and in some Gram-positives, and is a tripeptide (cystine, glutamate and glycine). Its two-step biosynthesis begins with fusion of cysteine to glutamate (catalyzed by the glutamate-cysteine ligase, GshA in *Brucella*) which produces γ-glutamylcysteine. This is the rate-limiting step, and is followed by the addition of glycine, catalyzed by the glutathione synthetase (GshB). Both steps are adenosine-triphosphate (ATP) dependent. GSH peroxidases can reduce peroxides to water by oxidizing two molecules of glutathione ($2\ GSH + H_2O_2 \rightarrow GSSG + 2\ H_2O$) as well as detoxify free radicals ($GSH + R\cdot \rightarrow 0.5\ GSSG + RH$). Oxidized glutathione (GSSG) is once more returned to its reduced state a NADPH-dependent GSH reductase (Gor) (123–125). Of note, not all bacteria synthesize glutathione (106, 126, 127). For instance, in *Mycobacteria*, the predominant LMW is mycothiol (MSH) and bacillithiol (BSH) is the most common LMW in *Bacillus* bacillithiol (128).

In the early endosome, transmembrane NADPH oxidase (NOX) will generate $O_2^{\bullet -}$ and H_2O_2 within the lumen of the vacuole. Furthermore, inducible nitric oxide synthases (iNOS) deliver RNS, such as $NO^{\bullet}$ and peroxynitrite $ONOO^{-}$. In some cell types, like neutrophils, myeloperoxidases can use the generated H_2O_2 to synthesize hypochlorous acid (HOCl), which is a two-electron oxidant that mainly targets proteins (106, 129, 130). Pathogens must thus find ways to survive the respiratory burst. Some employ numerous strategies to avoid the phagosome

altogether: for instance *Yersinia pestis* can disrupt signals that lead to the formation of the phagosome. Others, like *Salmonella enterica,* inhibit the production of ROS and RNS from within the phagocytic vacuole (131). When they do get bombarded by oxidants, pathogens can then employ various countermeasures to survive the assault, mainly by sensing the attack and increasing the basal levels of antioxidants they normally produce. Three master sensors are OxyR, PerR and OhrR. OxyR is a transcriptional regulator belonging to the LuxR family, that was initially identified in *Salmonella*. It is conserved in Gram-negative and is present in some Gram-positive bacteria. It is mainly a peroxide sensor, thanks to a conserved cysteine residue that is oxidized in response to hydrogen peroxide (it can detect down to 1 μM of extracellular H_2O_2). Upon activation, it upregulates transcription of catalases and SOD, among other genes, in response to peroxide stress. In pathogens, its depletion leads to attenuated virulence. PerR also mostly responds to H_2O_2. It is a transcriptional regulator of the Fur family that was discovered in *Bacillus subtilis*, and is mainly found in –but not limited to– Gram-positive bacteria. It is a zinc metalloenzyme that contains either Fe^{2+} or Mn^{2+} coordinated to histidine residues. Upon oxidation of the metal ion, repression of target genes is relieved in the case of PerR-Zn-Fe enzymes. In the case of PerR-Zn-Mn, induction does not de-repress the PerR regulon, which includes catalases and other genes to defend against oxidative stress and metal homeostasis genes. OhrR is also a transcriptional repressor, belonging to the MarR family, but is only minimally activated by H_2O_2. Instead, it is best inactivated by organic hydroperoxides (R-OOH) and sodium hypochlorite (NaOCl), that interact with a conserved cysteine residue, which in turn relieves repression of target genes, such as glutathione peroxidase, as well as genes involved in quorum sensing and tyrosine metabolism (132–134).

Pathogens that are unable to withstand or fight off oxidative stress are attenuated. For instance, the (p)ppGpp defective mutant in *Staphylococcus aureus* showed increase expression of PerR, among other oxidative stress regulators, during stationary phase, and was more sensitive to hydrogen peroxide stress as well as ciprofloxacin- and tetracycline-mediated killing (135). In *Brucella abortus*, SOD-deficient strains were found to be attenuated in both tissue culture and mouse infection models and pathogens developing resistance to RNS are a growing concern (136, 137).

1.6.3. Sulfur metabolism and pathogens

It is well established that sulfur metabolism plays a significant role in oxidant detoxification. Sulfur is essential for life and is incorporated into amino acids (cysteine and metabolism) as well as in essential cofactors (like coenzyme A (CoA), GSH, iron-sulfur clusters, and biotin). Different sulfur species can be imported into the bacterial cell. In the case of pathogens, inorganic sulfur is usually acquired as sulfate (SO_4^{2-}) and reduced first to sulfite (SO_3^{2-}) then to sulfide (H_2S), where it can be attached to organic molecules, such as O-acetylserine (OAS) to produce cysteine, O-succinyl-homocysteine to produce L-homocysteine, and onto formaldehyde to produce methanethiol. If the bacteria encode a particular OASS, CysM, thiosulfate can also suffice as the sulfur source. Organic sulfur sources are also often extensively used by bacterial pathogens, such as cysteine, methionine and glutathione. Cysteine is a precursor in the biosynthesis of many sulfur-containing molecules, including GSH (**see above**) and methionine may also provide a reliable sulfur source if converted into cysteine (**see below**). Furthermore, glutathione may be metabolized by bacterial cells to indirectly provide S. Indeed, GSH is produced by eukaryotic cells (from 1-10 mM intracellularly and close to 40 μM in human plasma), and synthesis is induced upon oxidative conditions to protect the host, for instance in the

case of the respiratory burst. Many pathogens encode a γ-glutamyl-transpeptidase (Ggt) that releases glutamate and γ-cysteinyl-glycine that can then release cysteine via peptidase action (138).

Cysteine and methionine are considered essential amino acids in mammals, as they need to be acquired by diet and are not synthesized. Cysteine biosynthesis from serine is a two-step pathway (see also **Chapter 3, Figure 6.16A**); the rate-limiting first step is catalyzed by an O-acetylserine transferase (known as SAT or CysE, encoded by *cysE* in bacteria). CysE attaches an acetyl group activated with coenzyme A (acetyl-CoA) to serine, forming O-acetylserine (OAS). Sulfide is then added to OAS by an acetylserine sulfhydrylase (OASS, encoded by *cysK* or *cysM* in bacteria) to form cysteine. The transsulfurylation pathway connects cysteine and methionine (see also **Chapter 4, Figure 6.30**). Cysteine may be converted to cystathionine by a cystathionine-γ-synthase (CGS, *metB*) then to homocysteine by a cystathionine-β-lyase (CBL, *metC*). This is known as the transsulfurylation pathway. Homocysteine is then converted to methionine by methionine synthases. The reverse pathway is also possible: methionine may be transformed to S-adenosylmethionine (SAM) by an SAM synthetase, then to homocysteine in a two-step reaction. From homocysteine, a cystathionine-β-synthase (CBS) catalyzes the formation of cystathionine which is then substrate for a cystathionine-γ-lyase (CGL) in the production of cysteine. These last two steps are part of what is known as the reverse-transsulfurylation pathway. Not all bacteria encode for the different enzymes involved in the transsulfurylation pathway (or the reverse process), thus modulating the different nutrients they can process and the different conditions they can grow in.

Defects in sulfur and cysteine metabolism can have effects on pathogen virulence. For instance, inhibitors targeting *Salmonella typhimurium* OASS have shown promising antimicrobial potential (139) and have also been tested in the treatment of *Mycobacterium tuberculosis* infection

(140). In *Acinetobacter baumanii*, functioning cysteine metabolism/sulfur assimilation pathways were found to be important in growth within the *Galleria mellonella* larvae-model of infection in a transposon-library experiment (141) and disruption of a cysteine metabolism transcriptional regulator (GigC) has been shown to play a role in virulence (142). In *Staphylococcus aureus*, defects in *cysK* have been shown to augment sensitivity to hydrogen peroxide stress and deletion of *tcyP* (encoding a transporter involved in sulfur acquisition) lead to competitive defects in murine models of infection (143, 144). Analysis of envelope proteins from *Brucella abortus* identified CysK1 (see below) as an immunogenic candidate when the envelope proteome was probed with antisera from a patient (145), and CysK2 from *Brucella abortus* was later shown to elicit an immunological response in mice (146).

1.7. *BRUCELLA OVIS*, THE RUNT OF THE LITTER

An unusual member of the *Brucella* genus is *Brucella ovis* (147). It is an understudied outlier within this almost-monomorphic family of pathogens. One of the most easily observable characteristic that sets *B. ovis* apart from other *Brucella* (and that it shares only with *B. canis*, as mentioned above) is the fact that it forms rough colonies on plates. This phenotype is linked to its lipopolysaccharide (LPS) structure. Indeed, the envelope of Gram-negative bacteria can be decorated with this large molecule, which is constituted of three main parts: 1) lipid A, a phosphorylated glucosamine disaccharide bound to fatty acids, anchors the LPS to the cell outer membrane; 2) the core saccharide component, which latches directly onto lipid A; 3) the O-chain, which is a repetitive glycan polymer, is the outermost portion of the LPS. Colonies of bacteria with intact LPS are smooth in appearance, while colonies with their LPS lacking the O-chain don't have a smooth colony texture, and are called "rough". LPS can come in many flavors and serve

various purposes, from granting structural integrity to the cell envelope, to protection, to aid in surface attachment. It is also a potent endotoxin, eliciting a strong immune response through toll-like receptor 4 (TLR4), thus promoting secretion of pro-inflammatory cytokines, nitric oxide and eicosanoids (148, 149), see **above**. Rough mutants of naturally smooth Brucellae lose virulence, but both *Brucella ovis* and *Brucella canis* remain naturally virulent, even though they lack the O-chain of their LPS (98, 150–152). Since LPS is involved in cell integrity, variation in envelope structure can render different *Brucella* species more or less susceptible to environmental stressors, specifically membrane stress (see the work from Herrou *et al.* (153) for examples of *Brucella* sensitivity to membrane stressors).

B. ovis is also the only non-zoonotic member of this family, as there are no reported cases of *B. ovis* infections in humans, and it appears to have a narrower tropism than other members of the genus: indeed, it is mostly limited to sheep, although there have been reports of cases within the red deer populations in New Zealand (154). This allows experimental procedures to only warrant BSL-2 facilities and practices, compared to the BSL-3 requirements for other *Brucella*. Furthermore, it seems to only enter the host through direct or venereal contact, and not through ingestion or inhalation. This limits its infective capacity drastically.

B. ovis also has a very specific tropism for the male genital tract in rams, where it causes epididymitis, orchitis and infertility. Compared to other *Brucella*, as mentioned above, *B. ovis* enters the mucosa, reaches the local lymph nodes then makes for the testes and epididymis, without affecting other organs. Indeed, for transmission to occur, the usual route needs an infected male to mate with a female and another healthy male to come along and mate with the ewe in the same estrus. Indeed, females tend to only suffer *B. ovis* infection transiently, and rarely experience abortion because of it (154–157).

Thus, there are many peculiarities concerning this outlier. At the genomic level, *Brucella ovis* shares *quasi*-genomic identity with other *Brucella* (94% nt identity with *Brucella abortus*), but has a highly degenerate genome. Indeed, about 244 of its genes are degenerated pseudogenes. These genes include enzymes involved in various metabolic and virulence processes. For instance, the urease pathway is non-functional in *Brucella ovis* (158). The urea-degradation pathway releases CO_2 and ammonia, and is utilized by various pathogens to survive the acidity of the stomach. This might explain why *B. ovis* cannot be transmitted via ingestion (80). Another example of an affected pathway is that of erythritol oxidation (159). This sugar is thought to be one of the preferred carbon sources for *Brucella*, is found in the placentas of goats, cows, and pigs, and has been linked to virulence in *Brucella* (160), possibly linking the strong tropism *Brucella* have for placental trophoblasts in animals and account for the limited capacity of *B. ovis* to induce abortion in ewes. Of note, a large deletion of 15 kb in the region that encodes for glycosyl transferases (which are in involved in LPS biosynthesis) accounts for the rough colony phenotype of *B. ovis* (80). Finally, this accumulation of pseudogenes and non-functional pathways also leads to its heightened sensitivity to partial CO_2 pressures (pCO_2) (161, 162), as discussed in detail in **Chapter 2** and **Chapter 4**.

Thus, *Brucella* lead a dangerous lifestyle (77, 82), and are constantly challenged by a dangerous environment (163). Studying how *Brucella* sense and react to these changing and challenging environments is key to understanding their physiology and infection, and eventually better fight this stealthy pathogen.

1.8. *BRUCELLA* CULTIVATION AND CO_2: A BRIEF HISTORY

As outlined above, *Brucella ovis* is an interesting outlier in the genus. One of its defining characteristics is its inability to grow in laboratory conditions without CO_2 supplementation. Carbon dioxide is an incredibly important molecule that permeates, directly or indirectly, into the metabolism, signaling, pH homeostasis, and in short, in the life of all organisms. Understanding how the environmental levels of CO_2 affect *Brucella ovis* remains an important question and is an old historical problem.

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 (53). 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 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 (53). In the years following the discovery of Bang and Stribolt, reliable axenic cultivation of *B. abortus* from clinical specimens remained a major challenge (164).

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 (165). 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 (166). 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 (167).

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 carbon into amino acids and nucleotides (168–170). Carbon-14 largely pooled in pyrimidines and in glycine: 90% of the carbon-14 detected in the amino acid pool was in glycine (170). Amino acids, purines, pyrimidines, various organic acids and cell hydrolysates were not able to substitute for CO_2 in their experiments (171). 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 (170). From these data, the authors concluded that fixation of CO_2 is a fundamental property of *B. abortus* strains.

1.8.1. 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 (52). This stands in contrast with *Brucella abortus* isolates from cattle that did not grow on agar surfaces (53) without CO_2 supplementation (165). Indeed, variation in atmospheric requirements for growth of these bacteria clouded the fact that these early investigators were actually studying closely-related species (172). Pioneering work by Alice Evans united the etiologic agents of Malta fever in humans and contagious abortion in cattle into a single group (173) that she classified under the genus *Brucella* (174). By 1930, it was established that certain *Brucella* isolates required CO_2 supplementation for axenic cultivation, while others could be cultivated without increased pCO_2 (175). 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 (166, 176). Buddle reported a similar phenomenon in an early description of *Brucella ovis* isolated from rams in New Zealand (177).

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 (166) 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 (166). 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 (178), supporting a model whereby this phenotype is likely a result of mutation at a single site.

1.9. CO_2 METABOLISM AND CARBOXYLASES

Upon noting the differences in requirement for CO_2 supplementation between different *Brucella* species, one might wonder why CO_2 is important in bacterial growth. Actually, carbon dioxide is a biologically important molecule at all levels, and it plays into various aspects of the life of all organisms. CO_2 is first and foremost the prime carbon source for autotrophs, and different organisms have evolved strategies to acquire carbon by fixing CO_2, though the relevance carbon dioxide acquisition is not limited to autotrophy. CO_2 fixation is dependent on the activity of carboxylases, and there are at least six different CO_2-fixation pathways that have been described. The importance of carboxylases is linked to the biological role of CO_2 in the growth and survival of all life. The first class of carboxylases are the autotrophic carboxylases fix CO_2, thus deriving biomass that all organisms use (directly or indirectly). The most prominent member of this class of carboxylases is, in fact, ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO). The second group of carboxylases are those involved in assimilatory pathways. Indeed, various

biological reactions require addition of a carboxyl group as a functional group to different substrates for additional transformations to occur, especially under anaerobic conditions. For instance, in aiding incomplete β-oxidation, where propionate assimilation requires propionyl-CoA to be transformed into (2S)-methylmalonyl-CoA by a biotin-dependent carboxylase since β-oxidation of fatty acids cannot metabolize the short odd-numbered fatty acid, allowing for the production of succinyl-CoA synthesis from (2S)-methylmalonyl-CoA, an intermediate of the tricarboxylic acid cycle (TCA cycle). Biosynthetic carboxylases belong to the third group, which includes carboxylases involved in the biosynthesis of specific types of molecules, for instance carboxylases involved in fatty acid metabolism. During the extension of the fatty acid chain, malonyl-CoA is obtained by addition of acetyl-CoA. A biotin-dependent carboxylase uses incorporated CO_2 to activate the acetyl-CoA molecule that will then be used to extend the chain. Carboxylases are also important in redox balancing and are involved in removing excess reducing equivalents (i.e. NAD(P)H). CO_2 thus functions as an electron sink, for instance in *Rhodobacter capsulatus*, a purple non-sulfur bacteria of the α-proteobacteria class, where RubisCO activity becomes essential if the bacteria are grown photoheterotrophically in the presence of highly reduced substrates like butyrate. Finally, the last group of carboxylases are those involved in anaplerotic reactions. When cells are actively growing and central carbon metabolism is active with many intermediates being drained from the TCA cycle as biosynthetic building blocks, the cycle needs its intermediates replenished to continue to function. Anaplerotic carboxylases like biotin-dependent pyruvate carboxylases and phosphoenolpyruvate carboxylases transform pyruvate or phosphoenolpyruvate, respectively, to oxaloacetate (179–181).

CO_2 is also the product of aerobic respiration and other biological reactions. It is present in the Earth's atmosphere at about 0.04% (though levels are increasing, https://www.co2.earth/)

but can be found in higher concentrations in different habitats, for instance within human cells and tissues. Indeed, mammalian cells have evolved up to 13 bicarbonate transporters to remove dissolved CO_2.

A lot of important carboxylation reactions require an essential cofactor: biotin. *Brucella* carries the whole set of genes required for biotin synthesis. Biotin, also known as vitamin B7, is a heterocyclic sulfur-containing cofactor. All biotin-dependent enzymes transfer CO_2 via a two-step reaction, which is also important in assimilation of carbon dioxide (see previous chapter). In the first step, biotin is carboxylated, binding CO_2. In the second step, the biotin cofactor releases the carbon dioxide which is either released or supplied to a substrate. There are three classes of biotin-dependent enzymes: carboxylases, decarboxylases and transcarboxylases. The latter (class III) is the only class that does not use free CO_2. Class II biotin-dependent enzymes free CO_2 from a substrate and release it. Finally, Class I enzymes are biotin-dependent carboxylases. They add free CO2 to an acceptor, which distinguishes between families of this class (182, 183).

1.10. CARBONIC ANHYDRASES

As detailed in **Chapter 2**, the molecular basis for the difference in in CO_2 requirement between *Brucella* lineages is the activity of a carbonic anhydrase. I will thus briefly introduce carbonic anhydrases and their biological relevance. As mentioned above, CO_2 in the atmosphere is not particularly abundant, and natural hydration to bicarbonate (which is the biologically relevant form of CO_2) occurs to slowly on a metabolic level. Carbonic anhydrases are enzymes that intervene in acquiring atmospheric (gaseous) CO_2 -which can freely pass through membranes- and trap it within cells in its hydrated ionic form, bicarbonate HCO_3^-.

Carbonic anhydrases were first purified 1933 from cow erythrocytes (184). They belong to a superfamily of metalloenzymes and are now known to present in all kingdoms of life. They catalyze the reversible hydration of carbon dioxide to bicarbonate ($CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$), although some have been shown to catalyze additional reactions, such as the hydration of carbonyl sulfide (COS) (185) and carbon disulfide (CS_2) (186) in bacteria, with important roles for their survival. Although different carbonic anhydrases display different kinetics, some are among the fastest known enzymes, $k_{cat}/K_M > 10^8$ M s^{-1} (187).

A divalent cation in the active site is required for catalytic activity, which is bound in a tetrahedral geometry, with one of the coordination bonds supplied by water. The most common is zinc (Zn^{2+}), which can be found lodged in the active site of all carbonic anhydrases, though the physiologic ion of choice varies between classes. To this day, eight distinct families have been identified (188): α-, β-, γ-, δ- , ζ-, η-, θ-, and ι-carbonic anhydrases. There is very little sequence and structural similarity between the different classes as carbonic anhydrases seem to have evolved independently and are a great example of convergent evolution. The enzymes from the different families differ in the preferred coordinated cation, in the 4D structure (i.e. some work as monomers, either as dimers, trimers, or higher) and in the specific organisms they may be isolated from.

The carbonic anhydrases that we will discuss in depth in **Chapter 2** and in **Chapter 4** are of the beta family. The beta family is found in Gram-negative and Gram-positive bacteria as well as in fungi, some Archaea, in chloroplasts of mono- and di-cotyledons, and in algae. They function as dimers or as multiples of dimers and also find Zn^{2+} in the active site. They can further be subdivided into class I or class II. In class I β-carbonic anhydrases, the active site is in an open conformation and shows only one histidine and two cysteines binding the metal ion. In class II,

the water molecule is replaced by an aspartate residue, which causes the active site to be in a closed conformation. These carbonic anhydrases are not active unless the pH increases above 8.3, where a conserved asparagine can then interact with the aspartate, freeing the active site and allowing a water molecule to enter (187–190)

1.10.1. Biological function of CO_2 and carbonic anhydrases

High CO_2 levels can damage to cells and tissues, and rapid conversion to bicarbonate allows for maintenance of pH homeostasis, ion secretion (191). Furthermore, carbonic anhydrases play an important function in biosynthetic pathways and in various anaplerotic reactions; for instance, they are involved in gluconeogenesis, lipogenesis (pyruvate carboxylase and acetyl CoA carboxylase), amino acid biosynthesis (through the activity of pyruvate carboxylase), and ureagenesis (192) (carbamyl phosphate synthetase I), pyrimide synthesis (carbamyl phosphate synthetase II). Then also play roles in CO_2 sensing and respiration (193) and in mammals they have been implicated in electrolyte secretion, bone resorption (194), calcification, in urinary acidification and bicarbonate reabsorption (in the kidneys) (195) and associated diseases (191). They also play an important role in tumorigenesis and tumor hypoxia. In algae and plants, carbonic anhydrases have also been shown to participate in photosynthesis and related biosynthetic reactions, and where, interestingly, CA expression can be induced by low CO_2 coupled to the presence of light (196–198). Furthermore, in diatoms, the delta and zeta families of CAs are involved in carbon fixation (199, 200) and η-carbonic anhydrases seem to play a role in *de novo* purine/pyrimidine biosynthesis, possibly presenting a novel target for anti-malarial agents (201–203).

Finally, carbonic anhydrases have been linked to virulence in diverse pathogens (187). In *Staphylococcus aureus* (a commensal and opportunistic pathogen Gram-positive pathogen of the

Firmicutes phylum), a recent report showed that MpsAB, a proposed sodium bicarbonate cotransporter, to be important in for growth in atmospheric CO_2 conditions as well as play a role in virulence (204). In the case of *Salmonella enterica* serovar Typhimurium (a member of the γ-proteobacteria and causative agent of typhoid fever), disruption of a putative carbonic anhydrase (mig-5) which is expressed within the host did not have an effect in an tissue culture infection assay, but was attenuated in a mouse model (205). In *Escherichia coli* (another commensal and opportunistic pathogen), the requirement for carbonic anhydrase activity was extensively studied examining the well-established *E. coli* carbonic anhydrases, Can and CynT (206). Lastly, *Brucella* appears to have three carbonic anhydrases, two β-carbonic anhydrase (BcaA and BcaB) and a γ-carbonic anhydrase (RicA). Studies on the role of these CAs' roles in metabolism, physiology, pathogenesis and inhibition have been conducted over the past eleven years and will also be discussed in further in **Chapter 2 and Chapter 4** in the context of the facultative intracellular pathogen *Brucella ovis* (161, 162, 207–210).

1.11. CONCLUSIONS

There are many shifts in the environment that bacteria need to deal with, in order to survive and thrive. The facultative intracellular pathogen lifestyle of *Brucella* is a fascinating model to query for this aspect. The challenge it faces within a host, specifically within a host cell, are as unique as the genetic determinants *Brucella* employs to withstand its natural habitat. Shifting it away from its natural niche and into a laboratory setting, the fastidiousness of *Brucella ovis* with regards to CO_2 is a bell that points to a particular form of adaptation of this understudied outlier within the *Brucella* family. We were curious to understand this phenotype more clearly and its link to other *Brucella* lineages (see **Chapter 2** and **Chapter 4**). Within the intracellular space, within

a vacuole, *B. ovis* is constantly stressed, persistently fighting off attacks of different natures. As such, understanding how it survives within the BCV, how this relates to stationary phase, and the metabolic requirements that underlie this line of defense is an interesting and pertinent avenue to explore, and our related findings are detailed in **Chapter 3** and further discussed in **Chapter 4**.

2. A CARBONIC ANHYDRASE PSEUDOGENE SENSITIZES SELECT *BRUCELLA* LINEAGES TO LOW CO_2 TENSION

2.1. PREFACE

The contents of this chapter were modified and adapted from its published form:

Varesio LM, Willett JW, Fiebig A, Crosson S., *Journal of Bacteriology* (2019)

DOI: 10.1128/JB.00509-19.

Data Sets may be found online.

2.2. INTRODUCTION

The requirement for CO_2 supplementation for *Brucella ovis* growth in laboratory conditions is an interesting feature of this member of the *Brucella* family. We sought to understand the genetic determinants that were responsible for this phenotype and link our discoveries to the *Brucella* family as a whole. Understanding how environmental levels of carbon dioxide affect growth of *Brucella ovis* and what advantage or disadvantage there may be in limited CO_2 assimilation is another angle at which to probe its biology and pathogenesis.

2.3. RESULTS

2.3.1. 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 (80). 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 **Figure 6.2A 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 7.1**).

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 nucleotide deletion within a six-nucleotide interval at the 3' end of a β-carbonic anhydrase gene (locus tag BOV_RS08635; $bcaA_{BOV}$) (**Figure 6.3A and Table 7.2**). 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 (**Figure 6.3A-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} (**Figure 6.3C**). Nucleotide deletions in $bcaA_{BOV}$ were the only mutations that were common to all mutant strains and missing from the parent (**Table 7.2**).

Similar $bcaA_{BOV}$ mutations in *B. ovis* strains that grow without CO_2 supplementation have been recently described by Perez-Etayo *et al*. (162), 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*. (211), 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 standard air after several days of incubation, we presume that wild-type *B. ovis* can replicate (albeit very slowly) in an unsupplemented atmosphere.

2.3.2. 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*) (**Figure 6.3D**). 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 (**Figure 6.2B and Figure 6.3E**). 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.

2.3.3. *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 7.3). 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.

2.3.4. 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}) (**Figure 6.3F**). 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) (**Figure 6.2C and Figure 6.3G)**. 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. 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 (**Figure 6.2C and Figure 6.3G**).

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 (**Figure 6.2D**). Thus, different single nucleotide deletions at the 3' end of *bcaA*$_{BOV}$ (**Figure 6.3C**) 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**).

2.3.5. 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 (187, 190) in diverse species(191, 212, 213), including *Brucella suis* 1330 (207, 208, 214). 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 (162) or bsCAII (214)) with the exception of one residue (**Figure 6.4A**). The *B. ovis bcaA1*$_{BOV}$ mutation thus results in a frameshifted gene product with a C-terminus that matches *B. suis* BcaA$_{BSU1330}$, as well as BcaA orthologs in other *Brucella* species, which we discuss below.

B. suis 1330 $BcaA_{BSU1330}$ is reported to be an active β-carbonic anhydrase (207, 214). 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) (206) was sufficient to support growth of the wild-type and Δ*bcaA B. ovis* strains in a standard atmosphere **(Figure 6.5A)**. Induction of *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 (**Figure 6.5B**), but enabled growth of wild-type and Δ*bcaA B. ovis* in an unsupplemented atmosphere (0.04% CO_2) (**Figure 6.5C**). We conclude that expression of *E. coli can* or *cynT* (**Figure 6.4B**) is sufficient to enable growth of WT and Δ*bcaA B. ovis* strains under the low CO_2 tension of a standard atmosphere and, therefore, that carbonic anhydrase activity in and of itself is sufficient to allow wild-type or Δ*bcaA B. ovis* to grow in 0.04% CO_2.

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

As outlined in the introduction, *B. ovis* and the majority of *B. abortus* strains require CO_2 supplementation for axenic cultivation (215). 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 (**Figure 6.6A**).

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 **(Figure 6.3)**, which

results in a corresponding frameshifted $BcaA_{BOV}$ C-terminus (**Figure 6.3 and Figure 6.6**). 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 **Figure 6.3A-C and Figure 6.6B,** 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 **Figure 6.6**), and *2)* the remainder encode full-length proteins (*B. abortus** in **Figure 6.6**) 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 (**Figure 6.7A, 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 contains sequences lacking the C339 insertion (52 out of 374 strains), resulting in a consensus BcaA gene product **(Figure 6.6 and Figure 6.7A, 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 frameshift site **(Figure 6.7A)**. 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.

2.3.7. 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* (**Figure 6.7B**). 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 6.7A**), and measured growth in a standard unsupplemented atmosphere (0.04% CO_2). Replacing $bcaA_{BOV}$ with $bcaA_{BAB2308}$ enabled growth of *B. ovis* in 0.04% CO_2 (**Figure 6.7C**), consistent with Perez-Etayo *et al*. (162). Replacing $bcaA_{BOV}$ with the frameshifted $bcaA_{BAB}$ pseudogene did not enable *B. ovis* growth in a standard atmosphere (**Figure 6.7C**). 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 (53, 165, 166, 176). Conversely, approximately one-quarter of sequenced *B. abortus* strains harbor a copy of *bcaA* in which the frameshift mutation at C339 has apparently been remediated **(Figure 6.7A)** 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. 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}, **Figure 6.7D**) 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 7.4)**. *B.*

abortus biovars 5 and 6 have been described as CO_2 independent (215) 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 (215), 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*. (215).

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.

2.3.8. 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 (162) 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 Chapter 1). 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 **Figure 6.8A** 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 (**Figure 6.9A**), and 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 Figure 6.9B)**. In contrast, shifting a *B. ovis* strain harboring the $bcaA1_{BOV}$ allele from 5% to 0.04% CO_2 had no effect on growth **(Figure 6.9C**) or gene expression using the same significance threshold **(Data Set 2 and Figure 6.8B)**. 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) (**Figure 6.8C and Figure 6.9D)**. 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 (**Figure 6.9E**).

2.3.9. 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 (**Figure 6.10)**. 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 Figure 6.10**). 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 (23, 29, 30, 216). In addition, multiple respiratory, energy metabolism, transport, and catabolic genes are downregulated upon CO_2 downshift **(Data Set 2, 3 and Figure 6.10)**. 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 6.9A)**.

Expression of genes implicated in response to stress, nutrient limitation, or stationary phase, including polyphosphate kinases (*ppk1*, *ppk2*) (217–219), small DUF1127-family proteins (*BOV_RS04430*, *BOV_RS13510, BOV_RS13790, BOV_RS13950*) (220), alternative sigma factors *rpoH* (221) and *ecfG/rpoE1* (222) 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 **(Figure 6.10 & Figure 6.11 and Data Set 2)**. Although *B. ovis* is urease negative and possesses a degraded urease gene cluster (80), transcription of the *ure* genes (and pseudogenes) is activated upon CO_2 downshift. Among the

most highly activated gene sets in low CO_2 are biotin biosynthesis genes **(Figure 6.10 & Figure 6.11 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 **(Figure 6.11 and Data Set 2)**.

2.3.10. 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 (223–225), 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 (226). 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 the *B. ovis* lineage, 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 (162), but we have shown that this gene is transcribed **(GEO accession GSE130678 and Figure 6.9E)** 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, both 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.* (80) (Table 7.5). *B. ovis* is urease negative (80, 227) and the *ure* gene clusters show features of degradation (80) 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 *ure* gene cluster (228). The T4SS provided us with a set of genes that are functional and intact in both species (229–231). The additional *B. ovis* pseudogenes included in our analysis are multiple erythritol catabolic genes and cytochrome oxidase genes, which have functional, full-length orthologs in *B. abortus* (Table 7.5). 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 7.3)**. Similarly, sequence differences are rare when comparing the *ure* genes or the T4SS genes within the 17 *B. ovis* strains. In short, the level of polymorphism we observe comparing 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 (**Figure 6.12**). *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 7.5) 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* (**Figure 6.12**), our analysis shows that many other pseudogene sequences besides $bcaA_{BOV}$ are also entirely conserved in across sequenced *B. ovis* isolates (Table 7.5).

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.

2.4. DISCUSSION

2.4.1. *Brucella* cultivation and the resurrection of a carbonic anhydrase pseudogene

During the course of transmission and infection, *Brucella* spp. must endure numerous shifts in the chemical state of the environment (59, 96), 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 (179). The importance of bicarbonate to life is evidenced

by the fact that at least 7 different families of carbonic anhydrases (188, 190, 191) 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 (162) 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*.

Almost all *Brucella* spp. can be cultivated without adding CO_2, but 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 (53, 165), and the knowledge that *B. abortus* strains require added CO_2 enabled early efforts to cultivate *B. ovis* from sheep tissue (65, 177). 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* (**Figure 6.3 & Figure 6.7and 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 (166, 177), 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 (232).

The fact that *B. ovis* and *B. abortus* lineages have independent insertion/frameshift mutations in *bcaA* that are apparently fixed in populations of each species suggests there may be

an adaptive advantage to loss of *bcaA* function. In the case of *B. ovis*, our data show that mutants harboring the functional *bcaA1*$_{BOV}$ allele grow significantly faster than wild-type at high partial CO_2 pressure (**Figure 6.3E and G**), which is what *Brucella* spp. encounter in mammalian host tissues. The spleen colonization phenotype of *B. ovis* strains carrying a functional *bcaA1*$_{BOV}$ allele does not differ from wild-type *B. ovis* in a mouse model of infection (162), but it is known that a slow growth rate can be advantageous for bacterial pathogens to maintain long-term infections in certain tissues (233, 234). Thus, it is possible that loss of *bcaA* function enhances *Brucella* fitness or persistence in tissues of natural (i.e. non-rodent) animal hosts by slowing growth. We note that *B. ovis* is almost entirely restricted to male reproductive tissue and is reported to be transmitted only through direct or venereal contact, via a transiently infected ewe (57, 157, 177). The rate of oxygen uptake and the partial pressure of CO_2 are higher in testes than in other tissues (235, 236), and it is therefore conceivable that *B. ovis* has evolved a lifestyle in which carbonic anhydrases are no longer required. *B. abortus* has a tropism for bovine reproductive tissue in males and females, and it may be also the case that loss of *bcaA* function influences its rate of growth and its persistence in these tissues, though this hypothesis remains to be tested.

There is no known *Brucella* reservoir outside of animals, but it is well established that presence of bacteria on aborted tissue is one of the main routes of transmission for *B. abortus* (77). A functional BcaA would seem to be advantageous for CO_2 assimilation when *B. abortus* is residing on such tissue, though the extraordinarily high titer of respiring brucellae on an aborted fetus (237) may lead to a high enough local pCO_2 that BcaA is unnecessary for growth in this context. Alternatively, *Brucella* growth may be arrested during the window in which it is outside its preferred host tissue. Growth impairment upon CO_2 limitation may, in fact, serve as an important signal for the bacterium during the transmission cycle. We do not currently understand

why loss of *bcaA* function is restricted to the *B. ovis* and *B. abortus* lineages, or if there is a particular feature(s) relating to CO_2 metabolism or transmission/infection biology that sets these species apart from other members of the genus. It is possible that all *Brucella* species, as intracellular animal pathogens, will eventually lose *bcaA* function.

2.4.2. CO_2 limitation elicits a starvation and virulence gene expression response

As outlined in the introduction, carbon-14 from supplied $^{14}CO_2$ was recovered primarily in pyrimidines and glycine in *B. abortus* (170, 238). 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 to the medium were unsuccessful (167). 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 (23, 30, 239). This response is evidenced by the downregulation of transcription and translation related genes as well as genes involved in respiration and central metabolism (**Figure 6.10**). 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.

2.5. 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 (**Figure 6.2A)**. 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 7.1**). 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 (240) (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/) (241). 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 **Figure 6.8**): 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. Samples were centrifuged at 4 °C (17000 × g) for 30min, 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 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. Two samples were extreme outliers: wild-type *B. ovis* (5% CO_2) replicate 3 and *B. ovis* $bcaA1_{BOV}$ (0.04% CO_2) replicate 1, 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 (242) (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 were removed if the number of the upregulated genes divided by the total number of genes assigned to that pathway was between 0.4 and 0.6.

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) (243). *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 the protocol of Wetmore *et al*. (243) and as previously described (153, 244) 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 7.4** 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 the 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. The heat map was built using Java Tree View version 1.1.6r4 (http://jtreeview.sourceforge.net).

2.6. SUMMARY

Brucella 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 CO_2 supplementation to cultivate particular species, which confounded early efforts to isolate *B. abortus* from diseased cattle. Differences in the capacity of *Brucella* species 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. Accordingly, we show that heterologous expression of functional *E. coli* carbonic anhydrases enables *B. ovis* growth in air. Growth of *B. ovis* is acutely sensitive to a reduction in CO_2 tension, while frame-rescued *B. ovis* mutants are insensitive to CO_2 shifts. *B. ovis* initiates a gene expression program upon CO_2 downshift that resembles the stringent response and results in transcriptional activation of its type IV secretion system. Our study provides evidence that loss-of-function insertion mutations in *bcaA* sensitize the response of *B. ovis* and *B. abortus* to reduced CO_2 tension relative to other *Brucella* lineages. CO_2-dependent starvation and virulence gene expression programs in these species may influence persistence or transmission in natural hosts.

3. *BRUCELLA OVIS* CYSTEINE BIOSYNTHESIS CONTRIBUTES TO PEROXIDE STRESS SURVIVAL AND FITNESS IN THE INTRACELLULAR NICHE

3.1. PREFACE

The contents of this chapter were modified and adapted from its published form:

Varesio LM, Fiebig A, Crosson S. *Infection and Immunity* (2021).

Data Sets may be found online.

3.2. INTRODUCTION

Brucella spp. are intracellular pathogens that have numerous mechanisms to contend with host-generated stressors and exploit host resources for growth. Within the host, brucellae are subject to nutrient limitation (102), phagosomal acidification (245), and direct attack from reactive oxygen and reactive nitrogen species (103) originating from the host-derived respiratory burst (104, 105). Dozens of genes involved in oxidative stress responses, acid stress responses, nutrient assimilation, and respiration have been implicated in the biology of *Brucella* infection (101). More recent studies have defined a role for the general stress response pathway in mitigation of multiple chemical stressors *in vitro* and in maintenance of chronic infection *in vivo* (222, 246). However, relatively little is known about the mechanisms *Brucella* spp. use to adapt to stresses encountered in axenic cultures during stationary phase. The study of stationary phase has the potential to inform

the discovery of genes that influence infection, intracellular replication, and survival (247) as there are postulated parallels between stationary phase physiology and the physiologic state of *Brucella* in the intracellular niche (102).

3.3. RESULTS

3.3.1. *B. ovis cysE* Tn-himar mutant strains have a fitness defect in stationary phase

We inoculated ≈1.5 × 10^9 *B. ovis* RB Tn-*himar* strains (see **Materials and methods** and (244)) into Brucella Broth in triplicate and collected samples at intervals throughout the growth curve: 0.05, 0.12, 0.9 and 2.4 OD_{600} (corresponding to early logarithmic, logarithmic, late logarithmic, and stationary phase). Barcodes were PCR amplified, sequenced, and tallied as previously described (243) to assess the relative abundance of each mutant strain in each sample. Our analysis yielded composite fitness scores for 2638 of 3391 annotated genes in *B. ovis* **(Data Set 1)**. Data for 118 mutants that exceeded a t-like test significance threshold ≥ 4 are presented in **Figure 6.13 (**see **Materials and methods)**. We observed the largest relative fitness score changes at OD_{600} = 2.4 (i.e., stationary phase) in this dataset.

To more rigorously assess mutants with fitness values that varied as a function of growth phase, we further filtered the genes to include only those with a fitness score ≥ |1| in at least one timepoint (see **Materials and methods**). We hierarchically clustered the 64 genes that passed this cutoff (**Figure 6.14A, Data Set 1**) and divided these clustered genes into four groups that displayed different fitness patterns throughout the growth curve (**Figure 6.14B**). Mutations in group 1 genes resulted in no fitness defect during exponential growth, but a fitness defect at OD_{600} = 2.4 (i.e. stationary phase). Genes in group 2 had negative fitness scores throughout the growth curve. These two groups contained the majority of mutants. Group 3 (four genes) had positive fitness scores in

log phase and a negative fitness score in stationary phase, while group 4 (three genes) had null or positive fitness scores at all phases of the growth curve. We clustered these genes by predicted functional category (**Figure 6.15A, Data Set 2**) and found that genes encoding purine metabolism enzymes and tRNA modification enzymes were enriched in group 1. However, the gene with the lowest fitness score in stationary phase, *BOV_RS06060* (old locus tag *BOV_1224*), is annotated as serine-O-acetyltransferase (*cysE*) (**Figure 6.14 and Figure 6.15B**). As such, we chose to further characterize the function of *cysE* in *B. ovis*.

3.3.2. *B. ovis* Δ*cysE* enters stationary phase prematurely and has reduced culture yield in vitro

B. ovis CysE has high sequence identity (52%) and similarity (73%) with the well-characterized CysE enzymes of *Escherichia coli* and *Salmonella enterica* (248), and is clearly classified as CysE in the NCBI conserved domain database (E-value = $4.3e^{-122}$, https://www.ncbi.nlm.nih.gov/Structure/cdd). This protein is therefore predicted to execute the initial step in cysteine biosynthesis, specifically the addition of an acetyl group from acetyl-CoA to serine, producing O-acetylserine (**Figure 6.16A**). To confirm the *cysE* stationary phase phenotype observed in the RB TnSeq experiment (**Figure 6.14**), we built a *B. ovis* strain harboring an in-frame deletion of *cysE* (Δ*cysE*). We grew Δ*cysE* in parallel with wild-type *B. ovis* ATCC 25840 (WT, **Figure 6.16B**) and observed that Δ*cysE* enters stationary phase earlier and terminates growth at a lower density than WT, thus corroborating the TnSeq result. This phenotype was complemented by the addition of 4 mM cysteine to the medium (**Figure 6.16B**) or by ectopic expression of the *cysE* gene from a *lac* promoter in the presence of IPTG (**Figure 6.16C**). Ectopic overexpression of *cysE* in a WT *B. ovis* background did not modify growth kinetics or the growth

curve shape compared to an empty vector control (WT/pSRK-EV) (**Figure 6.16C**). We conclude that *cysE* and cysteine biosynthesis are necessary for normal *B. ovis* growth yield in Brucella Broth.

3.3.3. *cysK1* and *cysK2* function redundantly in cysteine biosynthesis

CysK catalyzes the step in cysteine biosynthesis subsequent to CysE, namely the elimination reaction in which the acetyl group on O-acetylserine is displaced by sulfide to form cysteine (248) (**Figure 6.16A**). Given the stationary phase phenotype of Δ*cysE*, and the fact that this defect was chemically complemented by cysteine, we expected that mutations in cysteine synthase (*cysK*) should phenocopy Δ*cysE*. However, strains with transposon insertions in locus *BOV_RS09280* (old locus tag *BOV_1893*), annotated *cysK* in the NCBI RefSeq database, grew like wild type (**Data Set 1**). Growth of a strain harboring an in-frame deletion of *BOV_RS09280* also grew the same as WT *B. ovis* in Brucella Broth (**Figure 6.17A**), confirming the TnSeq result.

We considered that the lack of an apparent growth defect in the Δ*cysK* strain may be due to the presence of other genes with CysK activity. A possible candidate for such a gene is locus *BOV_RS05050* (old locus tag *BOV_1018*), which encodes a protein with 37% sequence identity and 53% similarity to *BOV_RS09280*. Hereafter, we refer to *BOV_RS09280* as *cysK1* and *BOV_RS05050* as *cysK2*. Strains with transposon insertions in *cysK2* also yielded a wild type phenotype in our TnSeq experiment (**Data Set 1**) and a strain harboring an in-frame deletion of *cysK2* (Δ*cysK2*) likewise grew the same as WT *B. ovis*. However, a Δ*cysK1* Δ*cysK2* double deletion strain exhibited a stationary phase/growth yield phenotype similar to Δ*cysE* (**Figure 6.17A**). Like Δ*cysE*, the Δ*cysK1* Δ*cysK2* phenotype was chemically complemented by the addition of cysteine to the medium (**Figure 6.17B**). The growth defect of the Δ*cysK1* Δ*cysK2* strain was genetically complemented by expressing either *cysK1* or *cysK2* from a plasmid (**Figure 6.17C**). We conclude that these two genes have redundant function in the cysteine biosynthesis pathway.

3.3.4. *B. ovis* Δ*cysE* and Δ*cysK1* Δ*cysK2* strains are sensitive to exogenous H_2O_2 stress

Cysteine is, of course, important for protein synthesis. It is also one of the three amino acids that comprise glutathione (GSH) (**Figure 6.18A**), which plays a central role in the mitigation of a variety of stressors in bacteria (124) including oxidative stress. We hypothesized that defects in cysteine biosynthesis would have consequences on GSH synthesis and sensitize cells to oxidative stress. We thus attempted to complement the Δ*cysE* growth phenotype by adding GSH to the medium (**Figure 6.18B**). GSH supplementation partially complemented the Δ*cysE* growth yield defect. GSH limitation may directly contribute to premature entry of *B. ovis* Δ*cysE* into stationary phase, or GSH addition may restore cysteine homeostasis upon GSH catabolism. Since GSH is known to be involved in decomposition of hydrogen peroxide to water (124) **(Figure 6.18A)**, we assessed whether Δ*cysE* was more sensitive to H_2O_2 stress. We grew WT and Δ*cysE* to stationary phase, washed the cells, treated them with H_2O_2 for one hour in phosphate buffered saline solution, and then enumerated CFU. The Δ*cysE* strain was ≈2000 times more sensitive to H_2O_2 than WT, and this sensitivity was rescued by either the addition of cysteine or glutathione to the medium during growth (**Figure 6.18C**). The hydrogen peroxide sensitivity phenotype of Δ*cysE* was genetically complemented by expression of *cysE* from the *lac* promoter of the pSRK plasmid (Δ*cysE*/pSRK-*cysE*) (**Figure 6.18D**). We further tested the sensitivity of the Δ*cysK1* Δ*cysK2* double deletion mutant to hydrogen peroxide treatment. Like Δ*cysE*, the Δ*cysK1* Δ*cysK2* strain was highly sensitive to peroxide treatment. The Δ*cysK1* Δ*cysK2* peroxide survival phenotype was chemically complemented by the addition of cysteine or glutathione to the medium (**Figure 6.18E**), and was genetically complemented by ectopic expression of either *cysK1* or *cysK2* (**Figure 6.18F**).

3.3.5. Δ*cysE* and Δ*cysK1* Δ*cysK2* have reduced viability in the intracellular niche

Brucella spp. primarily reside inside mammalian host cells. There are many challenges to growth and survival in the intracellular niche including nutrient limitation and exposure to stressors such as reactive oxygen species (ROS) (101). Given the *in vitro* growth and hydrogen peroxide sensitivity phenotypes of the *cys* mutants, we tested whether fitness of the Δ*cysE* and Δ*cysK1* Δ*cysK2* strains was compromised in the intracellular niche. We infected a human monocytic cell line, THP-1, that we differentiated into macrophage-like cells. Although entry (2 hrs post-infection, p.i.) to the macrophage was unaffected by deletion of *cysE* or *cysK1* and *cysK2*, there was a significant loss in recoverable colony forming units (CFU) of the *cys* mutant strains by 24 hrs p.i. relative to WT (**Figure 6.19A and B**). Thus *B. ovis* Δ*cysE* and Δ*cysK1* Δ*cysK2* enter host cells like WT but their fitness is compromised after entry.

In an effort to distinguish the relative contributions of intracellular killing and cysteine (nutritional) limitation on reduced fitness of the *cys* mutants, we enumerated CFU recovered from THP-1 cells at timepoints between 2 and 24 hrs p.i. Wild type and Δ*cysE* exhibit identical CFU loss between 2 and 8 hrs post-infection. WT begins to replicate by 12 hrs, but Δ*cysE* CFUs continue to decrease up to 24 hrs **(Figure 6.19C)**. The rate of recoverable CFU increase between 24 hrs and 48 hrs p.i. is similar between WT and the *cys* mutants providing evidence that *B. ovis* has access to cysteine in this environment **(Figure 6.19A and B)**. An attempt to interrogate a later time point (72 hrs) was confounded by egressing bacteria and subsequent killing by gentamicin in the tissue culture medium (**Figure 6.20)**. The intracellular infection defect we observe was complemented by expression of *cysE* from a plasmid (**Figure 6.21A**). We attribute partial complementation to the fact that *cysE* was expressed from a *lac* promoter on a replicating plasmid; there are challenges with full induction of transgenes from heterologous promoters in an intracellular infection context.

Given the ability of *Brucella* to infect multiple mammalian cell types, we next tested whether the *in vitro* infection phenotype of the *cys* mutants was particular to macrophages. We infected a sheep testis epithelial cell line (OA3.ts) (249), which is derived from a host tissue type that is relevant to *B. ovis* infection. OA3.ts entry was unaffected by the lack of *cysE*, but recovered CFU were significantly lower for Δ*cysE* by 24 hr p.i. (**Figure 6.19D**). This phenotype was partially complemented by ectopic expression of *cysE* from a plasmid (**Figure 6.21B**). The magnitude of the Δ*cysE* defect 24 hrs p.i. was greater in THP-1 macrophages than in the OA3.ts epithelial line (about 64-fold vs 16-fold, respectively; **Figure 6.22**). These *in vitro* infection data provide evidence that an intact cysteine metabolism system promotes *B. ovis* fitness in intracellular niche of multiple mammalian cell types.

3.4. DISCUSSION

3.4.1. A genome-scale search for *B. ovis* stationary phase mutants leads to cysteine metabolism

B. ovis is a widespread ovine pathogen that remains an understudied member of the *Brucella* genus. Using a RB TnSeq approach, we sought to identify genes that are important for *B. ovis* growth and/or survival in the late phase of axenic broth culture (i.e. stationary phase), with a larger goal of uncovering genes that are important for fitness in the intracellular environment. We identified multiple genes for which Tn-*himar* disruption resulted in reduced fitness in stationary phase. Among the expected mutants in this dataset is *rsh* **(***BOV_RS03230*), which controls the stringent response (25). Additionally, genes involved in purine metabolism, including *purF* have reduced fitness in dense culture. In *Mycobacterium smegmatis*, PurF influences survival during stationary phase (250), and purine metabolism is known to be important for growth of multiple

microbes in the intracellular and extracellular environments (251, 252). Multiple genes with a predicted role in tRNA modification also had diminished fitness in stationary phase. Transfer RNA modification enzymes have roles in translation quality control and can function to direct translation of specific transcripts under particular growth conditions (253). Given the phenotypes of tRNA modification mutants in stationary phase, it may be the case that these genes play a role of regulation of *Brucella ovis* physiology in the intracellular niche.

Tn-*himar* strains with insertions in *cysE* had the most diminished fitness in stationary phase, and *cysE* was therefore selected for follow-up studies. Sulfur and cysteine metabolism are central to microbial growth, and have been well studied in numerous pathogens (138). In *Brucella* spp., our understanding of sulfur metabolism is relatively limited though biosynthesis of sulfur-containing amino acids - cysteine and methionine - has been implicated in *Brucella melitensis* 16M infection of mice (254). Based on the high level of sequence identity/similarity to well-studied CysE enzymes and established structural data on *B. abortus* CysE (255), *B. ovis* CysE is presumed to catalyze biosynthesis of O-acetylserine from acetyl-CoA and serine. The subsequent step in biosynthesis of cysteine from O-acetylserine requires displacement of the acetyl group by sulfide, a reaction that is catalyzed by CysK in many bacteria. Our growth data clearly implicate *cysE* in the cysteine biosynthesis pathway, as the *in vitro* growth defect of Δ*cysE* is rescued by the addition of cysteine. These results support published data that *Brucella* spp. can assimilate cysteine as an exogenous organic sulfur source (244, 256).

Surprisingly, the growth phenotypes of strains with Tn-*himar* insertions in the gene annotated as *cysK* in the RefSeq database did not differ from wild type, which suggested redundancy at this biosynthetic step. Consistent with this hypothesis, we have presented genetic evidence that two related enzymes, CysK1 and CysK2, function redundantly to produce cysteine.

It is possible that CysK1 and CysK2 do not catalyze the same reaction, but rather determine cysteine biosynthesis through two distinct routes. A recent report of such a case is the cystathionine β-synthase of *Helicobacter pylori*, which retains some O-acetylserine sulfhydrylase activity (257); this enzyme shares primary structure features with *B. ovis* CysK2. CysK-family enzymes can also have functions beyond direct involvement in cysteine metabolism (258), which may influence interpretation of our results. Notably, *B. abortus* CysE (serine O-acetyltransferase) and CysK2 do not form a cysteine synthase complex (CSC) *in vitro* (259). This supports a model in which CysK2 participates in cysteine synthesis via a mechanism that differs from that catalyzed by the typical CysE-CysK CSC. We postulate that CysK1, rather than CysK2, binds to CysE to form the CSC in *Brucella*. The development of a defined medium that supports the growth of *B. ovis* would greatly facilitate future study of CysK1 and CysK2 functions in cells. Exploration of the possible intracellular fitness advantage gained by redundancy at the CysK step of cysteine biosynthesis is an interesting area of future investigation.

3.4.2. Cysteine, glutathione and hydrogen peroxide stress

The growth and peroxide survival defects of Δ*cysE* were partially rescued by addition of cysteine or glutathione to the medium. Though elevated intracellular cysteine enhances susceptibility to hydrogen peroxide stress in *Escherichia coli* (126, 260), we do not observe peroxide sensitization of WT or Δ*cysE B. ovis* upon addition of 4 mM cysteine. 4 mM cysteine was consistently more protective than 4 mM GSH in our assay. GSH is an important redox control molecule, but the protective effect of GSH supplementation against H_2O_2 may be indirect. Specifically, it's possible that *B. ovis* transports and metabolizes some of the GSH to release cysteine, which is one of the three component amino acids of GSH. *B. ovis* is predicted to encode a γ-glutamylcyclotransferase (*BOV_RS09395*), which catalyzes the cleavage of GSH to form

pyroglutamic acid and L-cysteinylglycine (123). The L-cysteinylglycine dipeptide could then be separated by peptidases to release cysteine. We nonetheless favor a model in which diminished GSH production (as a result of abolished cysteine production) in Δ*cysE* directly affects H_2O_2 detoxification and growth yield of Δ*cysE*. Glutathione metabolism is important in *B. ovis*: the *gshA* biosynthesis gene is essential based on our previously published TnSeq dataset (161). Moreover, Tn-*himar* insertions in *BOV_RS04850* (old locus tag, *BOV_0978*), which is predicted to encode a glutathione-disulfide reductase (*gor*) – that reduces GSSG to GSH – resulted in a significant fitness disadvantage throughout the growth curve .

3.4.3. Cysteine and growth and the intracellular niche

Our study provides evidence that cysteine biosynthesis contributes to *B. ovis* fitness inside mammalian host cells. Strains harboring deletions of *cysE* or both *cysK1* and *cysK2* were not defective in host cell entry but had significantly reduced recoverable CFUs at 24 and 48 hrs post infection in human macrophage-like and ovine epithelial cell lines. Reduced recoverable CFU of the *cys* mutants at 24 and 48 hours can be attributed to defects that are manifested between 2 and 24 hours post-infection. It is difficult to fully discern the relative contributions of nutritional restriction and enhanced oxidative stress sensitivity to attenuation of Δ*cysE* and Δ*cysK1* Δ*cysK2* *in vitro*. The similar observed rate at which recoverable CFUs increase between 24 and 48 hours suggests that cysteine levels are not limiting – at least after 24 hours. The ER-derived replicative *Brucella* containing vacuole (rBCV) supports bacterial replication (96) and can be established as early as 12 hrs p.i. (97); based on our data, we conclude that this compartment contains enough cysteine or cysteine-containing peptides to support growth. Of note, the defect of the Δ*cysE* strain was more pronounced in macrophage-like cells than in an ovine testis epithelial line. Sensitivity of Δ*cysE* and Δ*cysK1* Δ*cysK2* to ROS may underlie this difference in fitness between cell lines as

macrophages typically have a more robust respiratory burst than epithelial cells (261). A significant decrease in recoverable CFUs in Δ*cysE* is evident by 4 hr post-infection. This is a timepoint before intracellular *Brucella* replication occurs (97), supporting a model in which increased sensitivity to host killing underlies the reduced fitness of Δ*cysE* and Δ*cysK1* Δ*cysK2* in host cells.

Previous *B. abortus* TnSeq studies by Sternon and colleagues (262) did not identify *cysE* as a gene that was important for infection of Raw 264.7 macrophages. We observed significantly reduced *B. ovis* Δ*cysE* CFU relative to WT in Raw 264.7 cells by 24 hrs **(Figure 6.23)**, but the Sternon *et al*. experiment and our experiment differ in several ways. Indeed, the importance of cysteine metabolism in intracellular growth and/or survival may vary between *B. ovis* and *B. abortus* and between mammalian cell lines. Nonetheless, our data clearly provide evidence that a cysteine anabolism pathway in *B. ovis* is important for growth, stress survival and fitness in the intracellular niche.

Cysteine and methionine metabolic pathways are attractive targets to combat various pathogens (138) because, in mammals, these amino acids must be acquired from diet. Thus, compounds that disrupt cysteine metabolism are not predicted to have direct negative effects on mammalian metabolism. In fact, O-acetylserine sulfhydrylase (OASS; i.e. *cysK*) inhibitors are under investigation as therapeutics for *Mycobacterium tuberculosis* infections (140). Our work shows that genetic disruption of cysteine biosynthesis leads to a significant defect in *B. ovis* fitness within host cells. This pathway is therefore a possible target for combating brucellosis.

3.5. MATERIALS AND METHODS

Bacterial strains and growth conditions

Brucella ovis was grown on Tryptic Soy Agar (TSA, Difco Laboratories) plates, supplemented with 5% sheep blood (Quad Five) or in Brucella Broth (BB, Difco Laboratories, dissolved in tap water) for liquid cultures. Cells were incubated at 37 °C with 5% CO_2 supplementation. Kanamycin (Kan) 50 μg/ml, sucrose (5% w/v) or isopropyl β-D-1-thiogalactopyranoside (IPTG, GoldBio) at 1 mM or 2 mM, were added when required.
Escherichia coli strains were grown in lysogeny broth (LB, Fisher Bioreagents) or on LB + 1.5% agar (Fisher Bioreagents) plates at 37 °C with Kan supplemented at a concentration of 50 μg/ml when required. *E. coli* WM3064 strain, used for conjugation, was grown in the presence of 300 μM diaminopimelic acid (DAP, Sigma-Aldrich), as it is a DAP auxotroph.

Plasmid and strain construction

Deletion plasmid construction

To build the deletion strains, fragments of approximately 500 bp upstream and downstream of target genes were amplified with KOD Xtreme Hot Start polymerase (Novagen). These fragments were built so that 9 bases at both the 5' and 3' ends of the gene were maintained, keeping the gene product in frame to minimize polar effects. Purified DNA from *Brucella ovis* ATCC 25840 was used as a template. Amplified fragments were gel purified (ThermoFisher Scientific) and assembled into the pNPTS138 suicide deletion vector (digested with HindIII and BamHI restriction enzymes, New England Biolabs) using Gibson assembly (New England Biolabs).

Complementation plasmid construction

To build plasmids for genetic complementation, *cysE*, *cysK1* or *cysK2* were PCR amplified from *B. ovis* ATCC 25840 with KOD Xtreme Hot Start polymerase, gel purified (ThermoFisher Scientific) and Gibson assembled into pSRK (263) that had been digested with NdeI and KpnI

restriction enzymes (New England Biolabs). *cysE*, *cysK1* or *cysK2* were cloned downstream of P_{lac} (lactose, IPTG inducible promoter).

Delivery of plasmids to B. ovis

Constructed plasmids were transformed into chemically competent *E. coli* Top10 strains for plasmid maintenance. All plasmid inserts were confirmed by PCR and Sanger sequencing, and plasmids were delivered to *B. ovis* by conjugation using *E. coli* WM3064 as a donor strain. For conjugation, WM3064 donor strains were mated with *B. ovis* strains and spotted on TSA blood plates plus DAP and incubated overnight at 37 °C in a 5% CO_2 atmosphere. Mating spots were spread on TSA blood plates plus Kan (without DAP) to select for *B. ovis* plasmid acquisition. When deleting genes using the pNPTS138 plasmid, merodiploid clones were inoculated in Brucella Broth overnight to allow for a second crossover event, then spread on TSA blood plates plus sucrose (5% w/v) for counterselection. Single colonies harboring gene deletions were identified by patching clones on TSA blood plates with or without Kan. The putative deleted locus was PCR amplified using gene-flanking primers in Kan-sensitive clones, and the PCR fragment was resolved by gel electrophoresis to test whether the gene had been deleted. For a complete list of strains, plasmids, and primers, please see **Data Set 3**.

Growth Curves

Cells were inoculated from ≈48hr-old TSA blood plates into BB at densities ranging from OD_{600} 0.08 to OD_{600} of 0.2. Growth was assessed spectrophotometrically measuring optical density at 600 nm (OD_{600}). Growth curves were conducted at least three independent times with two or three technical replicated in each experiment. Representative curves are shown for each set of

strains. Where indicated, cysteine (4 mM), GSH (4 mM), Kan (50 μg/ml), or IPTG (1 mM) were supplemented upon start of growth experiment to the liquid media.

H_2O_2 survival assays

Cells were grown overnight in BB to stationary phase (OD_{600} of ≈2). Cells were pelleted and resuspended in Phosphate-Buffered Saline (PBS, Sigma) to achieve an OD_{600} of 0.15. 200 μl of cells were added to 1.8 ml of PBS or PBS supplemented with fresh H_2O_2 (15 or 20 mM final concentration), bringing the final OD_{600} to 0.015. Cells were then incubated at 37 °C in 5% CO_2 for 1 hr before spotting aliquots of a 10-fold serial dilution series on TSA blood plates. CFUs were enumerated after 48 hrs incubation at 37 °C in 5% CO_2. Experiments were repeated at least three times with each sample in duplicate or triplicate in each experiment.

DNA extractions

Cells from 1 ml of stationary phase culture were pelleted by centrifugation, washed once in PBS, and resuspended in 100 μl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) supplemented with 1 μg/ml RNAseA. Cells were lysed by addition of 0.5 ml GES lysis solution (5 M guanidinium thiocyanate, 0.5 M EDTA pH 8.0, 0.5% v/v Sarkosyl) and 15 min incubation at 60 °C. 0.25 ml cold 7.5 M ammonium acetate (Fisher Bioreagents) was added, and mixture was incubated on ice for 10 min. 0.5 ml of chloroform (Fisher Bioreagents) was added to separate the DNA, samples were vortexed and centrifuged. Aqueous top phase was moved to a fresh 1.5 ml centrifuge tube and 0.54 volumes of cold isopropanol was added to precipitate the DNA. After centrifugation, isopropanol was discarded and pellets were washed three times in 70% ethanol before resuspending pellets in TE buffer. Concentration and purity of the extracted DNA were determined spectrophotometrically (NanoDrop One, Thermo Scientific).

Barcoded TnSeq

A *B. ovis* RB Tn-*himar* library was built and mapped as described (244). Briefly, *E. coli* APA752 (a WM3064 donor strain carrying a pKMW3 *mariner* transposon library) was conjugated into *B. ovis bcaA1* (161) under atmospheric CO_2 conditions. Kan resistant colonies were collected, grown to OD_{600} = 0.6 and frozen in 1 ml aliquots. Genomic DNA was extracted and the Tn insertion sites mapped as previously described (243).

To identify genes that confer a fitness advantage in stationary phase, the *B. ovis* Tn-*himar* library was inoculated in BB in a 5% CO_2 environment in triplicate at an OD_{600} of 0.0025. An aliquot of each initial culture was collected as the reference time point. Cultures were then grown to stationary phase, with samples harvested throughout the growth curve at OD_{600} of 0.05, 0.12, 0.9 and 2.4. Cells from each sample were pelleted by centrifugation and resuspended in water. Barcodes were amplified from approximately 1.5×10^8 cells per PCR reaction (see **Table 7.6**) with primers that both amplified the barcodes and added indexed adaptors (243). Amplified barcodes were then pooled, purified and sequenced on an Illumina HiSeq 4000.

Fitness scores for each gene were calculated following the protocol of Wetmore and colleagues (243) using scripts available at https://bitbucket.org/berkeleylab/feba. Genes for which 2 out of 3 samples had at least one timepoint with a t-like statistical significance score (243) $\geq |4|$ were included in subsequent analyses. A heat map of fitness scores of genes passing this filter is shown in **Figure 6.13** and the raw fitness data are in **Data Set 1**.

Finally, we averaged the fitness values of the three replicates and kept mutants that had an average fitness score $\geq |1|$ in at least one time point. Mutants in this group with a standard deviation ≥ 0.75, were manually inspected and extreme outlier points were removed from a total of six genes.

The genes with adjusted average and standard deviation values are shown in red in **Data Set 1**. Heat map of averaged fitness values is shown in **Figure 6.14**.

Tissue culture

All tissue culture cells were grown at 37 °C with 5% CO_2 supplementation. THP-1 cells (ATCC TIB-202) were cultured in Roswell Park Memoriam Institute medium (RPMI 1640, Gibco) + 10% Fetal Bovine Serum (FBS, Fisher Scientific). The RAW 264.6 (ATCC TIB-71) and the OA3.ts (ATCC CRL-6546) cells were grown in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% FBS.

Infection assays

THP-1 cells were seeded at a concentration of 10^5 cells/well in 96-well plates and phorbol myristate acetate (PMA) at final concentration of 50 ng/μl was added to induce differentiation into macrophage-like cells for 48-96 hrs prior to infection. OA3.ts cells were seeded at a density of 5 × 10^4 cells/well in 96-well plates for 24 hrs prior to infection. *B. ovis* cells were resuspended from a 48 hr old plate in RPMI + 10% FBS or DMEM + 10% FBS and added to tissue culture plates on the day of infection at multiplicity of infection (MOI) of 100 for THP-1, and at an MOI of 1000 for OA3.ts cells. When infecting with the complementation strains carrying the pSRK plasmid, the *Brucella* strains were struck on TSA blood plates ith Kan and IPTG 48 hrs prior to infection, and 2 mM IPTG was added to the tissue culture media throughout the duration of the experiment. Plates were spun for 5 min at 150 ×g and incubated for 1 hr at 37 °C in 5% CO_2. Fresh media was supplied containing 50 μg/ml of gentamicin and incubated for another hour. Cells were then washed once with PBS and once in H_2O and then lysed with H_2O for 10 min RT at 2 hrs, 24 hrs, and 48 hrs post infection. Lysates were serially diluted, spotted on TSA blood plates and incubated

at 37 °C in 5% CO_2 for 48 hrs to enumerate CFUs. Experiments were repeated at least 3 times with three technical replicates.

3.6. SUMMARY

Brucella ovis is an ovine intracellular pathogen with tropism for the male genital tract. To establish and maintain infection, *B. ovis* must survive stressful conditions inside host cells, including low pH, nutrient limitation, and reactive oxygen species. These same conditions are often encountered in axenic cultures during stationary phase. Studies of stationary phase may thus inform understanding of *Brucella* infection biology, yet the genes and pathways that are important in *Brucella* stationary phase physiology remain poorly defined. We measured fitness of a barcoded pool of *B. ovis* Tn-*himar* mutants as a function of growth phase and identified *cysE* as a determinant of fitness in stationary phase. CysE catalyzes the first step in cysteine biosynthesis from serine, and we provide genetic evidence that two related enzymes, CysK1 and CysK2, function redundantly to catalyze cysteine synthesis at steps downstream of CysE. Deleting either *cysE* (Δ*cysE*) or both *cysK1* and *cysK2* (Δ*cysK1* Δ*cysK2*) results in premature entry into stationary phase, reduced culture yield and sensitivity to exogenous hydrogen peroxide. These phenotypes can be chemically complemented by cysteine or glutathione. Δ*cysE* and Δ*cysK1* Δ*cysK2* strains have no defect in host cell entry *in vitro* but have significantly diminished intracellular fitness between 2 and 24 hours post infection. Our study has uncovered unexpected redundancy at the CysK step of cysteine biosynthesis in *B. ovis*, and demonstrates that cysteine anabolism is a determinant of peroxide stress survival and fitness in the intracellular niche.

4. DISCUSSION, CHALLENGES FOR THE FUTURE, AND CONCLUSIONS

Our work in *Brucella ovis* in the context of different environmental stimuli and its capacity to interact and react to these factors has increased our understanding of this furtive bacterium. In **Chapters II and III** I detailed our discoveries in relation to CO_2 sensitivity and the implications of cysteine metabolism within the host intracellular environment. In this final chapter, I will briefly discuss some challenges and possible future directions.

4.1. CO_2 SENSITIVITY

In **Chapter 2**, I addressed a historical observation about CO_2-dependence of different *Brucella* species and biovars. In trying to understand the genetic determinants that underlie *B. ovis* sensitivity to CO_2 levels, we uncovered a complex collection of alleles within different *Brucella* lineages. We show that these alleles determine *Brucella* capacity to grow in atmospheric conditions. Furthermore, when the allele in question encodes for a non-functional carbonic anhydrase, this sensitizes the strain to CO_2 levels, with implications at the transcriptional level for a multitude of genes, including those affecting virulence.

The increased sensitivity in *Brucella ovis* to partial CO_2 pressures is an interesting discovery. Indeed, the $bcaA_{BOV}$ pseudogene may be under neutral selection and in the process of being lost within the *Brucella* population, with *Brucella ovis* and most *B. abortus* biovars as the forerunners of this process. There is no evident morphological difference between WT *B. ovis* and *B. ovis bcaA1* strains grown in 5% CO_2 (**Figure 6.24**), as well as no difference at the transcriptional

level (**see Figure 6.8 and Figure 6.9**). The lack of an obvious *in vivo* advantage (162) to harboring the non-functioning carbonic anhydrase pseudogene provides evidence that this model is accurate. Of note, García-Lobo and colleagues did discern a slight fitness advantage for CO_2-dependent *Brucella abortus* 292 strain versus a derived CO_2-independent mutant (264). This experiment was conducted in female BALB/c mice, where spleen and lungs were tested for *Brucella* colonization by CFU plating after 8 weeks of infection. To distinguish between the two strains, dilution plates were either grown in the presence or absence of CO_2 supplementation. While these results are interesting and offer a potential *in vivo* advantage to harboring the non-functional *bcaA* pseudogene allele, it should be noted that there are a couple of points to this experiment that should be addressed and perhaps more fully explored. On one hand, neutral selection may be causing an increase of the ratio of CO_2-dependent bacteria (from the CO_2-independent population) once the inocula is within the host (and in a high CO_2 environment), so perhaps using a different selection method to distinguish between the two inoculated strains may be more apt, like different antibiotic resistance genes stably transformed within the strains. Furthermore, the *bcaA* pseudogene (and consequently CO_2-dependence) is found in all *B. ovis* isolates (to date) and in most *B. abortus* biovars, but there is no evidence of other *Brucella* species harboring this non-functional allele. So, if indeed there is an advantage to losing carbonic anhydrase function, it might be linked to these two *Brucella* species. Since *Brucella ovis* and *Brucella abortus* do not share preferred host species, LPS structure, nor are the closest kin based on 16S phylogenetic analysis, it is unlikely that a murine infection model will serve well as a basis to understand the cause and physiology of this loss-of-function selection process. Indeed, it is possible that a set of genes or gene functions that are shared between *B. abortus* and *B. ovis*, or some *B. abortus* biovars, determine an advantage to losing carbonic anhydrase function. If this were the case, is there residual function or a novel

function for the *bcaA* pseudogene? Again, that may be the case if the murine competitive infection assays are telling of what is happening in the environment during naturally-occurring *Brucella* infections. This might indeed be the reason we do not see more mutations accumulate within $bcaA_{BOV}$ (as shown in **Figure 6.12**). Further studies, for instance at a population level, where freshly isolated strains were immediately screened for CO_2 dependence and their *bcaA* alleles sequenced, might help understand this phenomenon better. Furthermore, the site of isolation might also be telling: for instance, whether *B. abortus* isolated from an aborted fetus carry more CO_2-independent strains, as the *Brucella* now are in an environment with only 0.04% CO_2, versus *Brucella abortus* isolated from a lymph node or the host testes.

4.2. OTHER *BRUCELLA* CARBONIC ANHYDRASES

Interestingly, *Brucella* do seem to encode for at least three carbonic anhydrases, two members of the β-family (BcaA and BcaB) and a putative γ-carbonic anhydrase (RicA) (209, 265, 266). RicA has never shown any carbonic anhydrase activity, and is one of the few identified effector proteins secreted by the T4SS (see Chapter 1**, The intracellular niche**). *BcaA* from *B. suis* 1330 has been biochemically characterized by Joseph *et al.* (214), and has modest CO_2 hydration activity, with a k_{cat} of 1.1×10^6 s^{-1}. I built a model structure for $BcaA1_{BOV}$ (**Figure 6.25A**) based on sequence homology to *Pisum sativum* (**Figure 6.25B and C)** with which it shares ~35% amino acid identity. The active site with the zinc bound to two cysteines and a histidine, along with the aspartic acid that acts as the fourth zinc ligand instead of a water molecule are shown in green in detail in **Figure 6.25A** right panel (characteristic of the type II β-carbonic anhydrase family, see Chapter 1 and also orange asterisks in alignment in **Figure 6.4B).** *E. coli* carbonic anhydrase, CynT, instead, shares 41% identity with the crystal structure of a lyase of

Synechocystis sp. (strain PCC 6803 / Kazusa), and is shown in **Figure 6.25B and D)**, and shares 33% identity to *Pisum sativum*. The portions in teal for the three structures indicate the area that is affected by the frameshift in *B. ovis*. The crystal structure of the second *E. coli* carbonic anhydrase, Can, is also shown (**Figure 6.25E)**. Based on the predicted structure (see also García-Lobo *et al.* (264)) of $BcaA1_{BOV}$, it seems plausible that the SNPs that transform BcaA from a functional carbonic anhydrase to a pseudogene in *Brucella ovis* do not affect the active site directly (**Figure 6.25, green**), but rather affect dimerization (**Figure 6.25, teal ribbons**). While these mutations do, overall, affect function of the resulting enzyme, it may be that the *bcaA* pseudogene is still capable of binding CO_2 or bicarbonate, which might in turn serve an ulterior function, such as sensing or sequestering these molecules. Further biochemical studies on this enzyme and crystallographic data might help better understand the nature of this protein. Of note, our attempts of purifying $BcaA_{BOV}$ were unsuccessful, as it was mostly recovered in the insoluble fraction (inclusion bodies, I.B.), and would rapidly precipitate out of solution (**Figure 6.26** and data not shown). This was not the case for *B. ovis* BcaA1 or the *Brucella suis* 1330 carbonic anhydrase $BcaA_{BSU}$, which was used as a positive control, as previous reports indicate successful purification (207, 214). This could suggest that $BcaA_{BOV}$ is indeed unable to dimerize, which is important for carbonic anhydrase function and enzyme stability. Finally, since carbonic anhydrases can also catalyze other hydration reactions other than CO_2 hydration (see Chapter 1), it is possible that the *bcaA* pseudogene retains (or gains) this additional function while losing carbonic anhydrase activity. Further biochemical studies would be necessary to fully understand what the frameshift mutation that inactivates $BcaA1_{BOV}$ does to this enzyme.

The other *Brucella* β-carbonic anhydrase, *bcaB*, was also purified in *B. suis* 1330 by Joseph and colleagues (207). It showed a more modest activity compared to *B. suis* 1330 *bcaA*, with a k_{cat}

of 6.4×10^5 s^{-1}. Interestingly, the *B. ovis bcaB* allele is missing 24 nucleotides in the middle of the gene (starting at codon 70) compared to other *Brucella* alleles (**Figure 6.27A**). There are two amino acid substitutions between BcaB in *Brucella suis* and *Brucella abortus* that seem to affect activity (162). In fact, *B. abortus* BcaB ($BcaB_{BAB}$), carries a glycine in position 76 instead of a valine (which is found in *B. suis* 1330 BcaB, $BcaB_{BSU}$). Perez-Etayo and colleagues have shown that two different *B. suis* strains (which are CO_2-independent) can still grow when lacking either β-carbonic anhydrase, that is $BcaA_{BSU}$ or $BcaB_{BSU}$. On the contrary, *B. abortus* CO_2-independent strain, *B. abortus* 2308W, cannot grow in atmospheric conditions if $bcaA_{BAB2308W}$ is deleted. The only common amino acid in the two *B. suis* BcaB alleles that is different in *B. abortus* is indeed the valine in position 76 (glycine in *B. abortus*), suggesting that the inactivity (or reduced activity) of $BcaB_{BAB2308W}$ is due to this amino acid substitution (162). Interestingly, this codon is not present in $BcaB_{BOV}$, as it is in the middle of the 8 amino acid-deleted portion (**Figure 6.27A**), indicating the importance of this locus for this carbonic anhydrase. We expressed $bcaB_{BAB2308}$ ectopically in WT *B. ovis* to attempt to rescue growth when CO_2 was not supplemented, but either $bcaB_{BAB2308}$ is not functional in *B. ovis* or it is not active enough to compensate for the *bcaA* pseudogene, because its expression upon IPTG induction did not allow *B. ovis* to grow in atmospheric CO_2 conditions (**Figure 6.27B and C**). Of note, Perez-Etayo et al. also found that $bcaB_{BSU1330}$ did not rescue growth of wild type CO_2-dependent *B. abortus* strains, which carried the $bcaA_{BAB}$ pseudogene (see **Figure 6.7**). This may suggest either that $BcaB_{BSU}$ activity is not sufficient to rescue growth in *B. abortus* when the *bcaA* allele is non-functional or that the functional and non-functional monomers are interacting in a way that either affects dimerization or yields a non-functional multimer, as this experiment was not conducted in a Δ*bcaA* background (162). As of now, the exact role for *bcaB* is in CO_2 fixation in these different *Brucella* species is not understood,

but further functional and structural characterization of this enzyme might lead to better understanding of the CO_2 acquisition network in *Brucella*, and shed light on the species-specific behaviors.

4.3. EFFECTS OF ALTERED SULFUR AND CYSTEINE METABOLISM IN *B. OVIS*

In **Chapter 3**, we found that the disruption of *cysE* strongly affected stationary phase in *Brucella ovis*. As *cysE* encodes for the enzyme responsible for the first committed step in *de novo* cysteine biosynthesis, this led us to investigate the role of sulfur metabolism in *Brucella ovis*. Specifically, we noted redundancy in the pathway at the level of the cysteine synthase step (CysK) and assessed how resistance to hydrogen peroxide is affected by interfering with cysteine biosynthesis. Furthermore, lack of cysteine production hinders the capacity of *B. ovis* to survive and replicate the hostile intracellular environment.

Sulfur is essential for life and it is therefore not immediately surprising that disruption of sulfur metabolism (or metabolic pathways that rely on sulfur availability, such as cysteine biosynthesis) would lead to important detrimental effects within *Brucella ovis*. What was surprising was that an important gene involved in *de novo* cysteine synthesis from serine (*cysE*) emerged as the costliest loss in stationary phase. Cysteine biosynthesis pathways interconnect in various ways in different organisms, as not all enzymes involved in different metabolic pathways are always present. Therefore, the specific effects that come with alteration of cysteine metabolism in *B. ovis* will have *B. ovis*-specific consequences, which helped us understand in more detail the nature of this pathogen's interaction with the challenging intracellular environment it encounters.

4.3.1. Coloration of Δ*cysE* strains

When scraping Δ*cysE* off of blood plates, we noticed an interesting occasional phenomenon: the pellet had a dark-purplish tint. Upon further attempts to investigate the phenomenon, we scraped pellets from different genetic backgrounds and compared the colors (**Figure 6.28A and B**). Interestingly, it seemed that the dark, purplish color matched strains in which *cysE* was not expressed. In fact, either WT *B. ovis* or *B. ovis*/pSRK-EV, *B. ovis* ΔcysE/pSRK-EV, and *B. ovis* Δc*ysE*/pSRK-*cysE* (in the absence of IPTG) always had a paler-to-white coloring, compared to *B. ovis* Δ*cysE* or the complementation strain grown on IPTG plates. Unfortunately, attempt to quantitatively describe this phenomenon failed, partially due to the transient nature of the color in the first place – cells would bleach rather quickly – and partially because of the difficulty in stably reproducing the phenomenon. We assessed that the time since thawing mattered, the chance of color being observed higher from strains struck 48 hours prior to collection, later times (when *B. ovis* colonies assume a more whitish coloring on the plates compared to the initial grayish hue) were never colored (**Figure 6.28A, bottom**). We did not determine a correlation with thickness of the lawn of *Brucella* (data not shown) nor with the presence of blood on the plate (**Figure 6.28A**). Nevertheless, it is an interesting phenotype that might be intriguing to pursue. Indeed, in the yeast pathogen *Cryptococcus neoformans*, Toh-e and colleagues were investigating sulfur amino acid metabolism and stumbled upon a similar observation: their Δ*cysE* or Δ*cysK* (*cys2*Δ and *cys1*Δ, respectively), when struck on plates with complex media supplemented with 1 mM copper sulfate ($CuSO_4$), acquired a rusty, dark-purple coloring, while wild-type *C. neoformans* remained white (267). This phenotype was not complemented by cysteine supplementation. The authors suggested the reason behind the coloring to revolve around H_2S accumulation when cysteine biosynthesis was halted because O-acetylserine

was not available as an acceptor for the reduced sulfur. We did not detect changes in coloring specific to the use of plates supplemented with $CuSO_4$ (data not shown).

4.3.2. Sensitivity to transition metals

In *C. neoformans*, it was also observed that interfering with cysteine biosynthesis affected sensitivity to heavy metals, such as copper ($CuSO_4$) and cadmium (in the form of $CdCl_2$). We first tested our Δ*cysE* strains with different concentrations of these metals by imbibing discs and analyzing the zone of clearance on a lawn of *Brucella ovis* (**Figure 6.28C**). We observed that in the presence of 10 mM cadmium chloride-imbued discs, there was a zone of clearing in *B. ovis* that was not clearly defined, while on lawns of *B. ovis* Δ*cysE*, a shiny metallic zone appeared **Figure 6.28C, white arrow)**, possibly suggesting metal precipitation. The zone of clearing round copper sulfate discs did appear slightly larger in strains lacking *cysE* compared to wild-type *B. ovis* (**Figure 6.28C**) and is in line with the observations made for *Cryptococcus neoformans* (267), though not as drastic. Of note, plates left in the incubator over time developed single defined colonies in the zone of inhibition of the $CuSO_4$, albeit close to the perimeter, in both WT and Δ*cysE* strains, suggesting the birth of copper-resistant mutants (**Figure 6.28C, white asterisks**). The origin of this shiny metallic zone in the presence of cadmium is also an interesting area of future investigation.

4.4. CYSTEINE AND METHIONINE METABOLISM IN *BRUCELLA OVIS*

4.4.1. *B. ovis* presumably does not encode for genes required for the reverse transsulfurylation pathway

Cysteine may be synthesized via a second metabolic pathway, from methionine, known as the reverse transsulfurylation pathway. Specifically, this pathway involves are a cystathionine beta-synthase (CBS, or *mccA*), which produces cystathionine from homoserine (a methionine derivative) and serine, and a cystathionine γ-lyase (CGL, or *mccB*), which yields cysteine (see Chapter 1). To attempt to test whether the reverse transsulfurylation pathway is functional in *B. ovis*, we attempted to rescue the Δ*cysE* phenotype by supplementing methionine in the media. Increasing amounts of methionine (up to 10 mM) did not restore growth of the *cysE*-null strain to WT levels (**Figure 6.29A**). It is possible that the reason for this observation is that methionine is not imported within the cell. We deem this unlikely, as we identified a putative methionine transporter in the *B. ovis* genome (*metQ*, locus tag *BOV_RS14915*) and given the promiscuity of amino acid transporters in general (268). We also did not identify homologs to CBS and CGL in *B. ovis*. Given these results, we assume that *Brucella ovis* does not have a functional reverse transsulfurylation pathway.

4.4.2. The redundancy of the two *cysK* genes in *B. ovis*

We found the functional redundancy of *cysK1* and *cysK2* in *B. ovis* with regards to cysteine biosynthesis intriguing. As mentioned in **Chapter 1**, cysteine biosynthesis can occur from serine (*de novo*) or via the reverse transsulfurylation pathway (**Figure 6.30**). We did not find any genes that could putatively encode for enzymes involved in the reverse transsulfurylation pathway. However, recent reports describe a novel peculiar cystathionine β-lyase which used OAS instead of serine to produce cysteine directly from homocysteine (O-acetylserine-dependent CBS, or OCBS) (269–271). In 2019 same kind of enzyme was described in the enteric pathogen *Helicobacter pylori* (257). Detailed analysis highlighted key amino acid portions that compare OASS, CBS and OCBS. We wondered whether one of the two *Brucella ovis* genes identified as

showing *cysK* function (*cysK1* and *cysK2*, **see4Chapter 3**), may actually encode for an OCBS. Upon comparison of five significant regions the authors identified in *H. pylori* as determinants of the specific enzyme type and function, it seems that *cysK1* looks more like an O-acetylserine sulfhydrylase (OASS), while *cysK2* is more similar to an OCBS (

Figure 6.31 **and ref**). However, since *B. ovis* does not appear to encode for CGL, the role for *cysK2* as an OCBS is obscure. More experiments would be necessary to understand the catalytic abilities of these important players, and a functioning minimal or determined media would aid future experiments greatly. To this end, testing homologous enzymes in *B. abortus* (for which a functional minimal media has been identified) might be a good starting point.

4.5. CONCLUSIONS

The role of sensitivity to environmental CO_2 levels in *B. ovis* has shed further light in understanding the physiology and evolution of this outlier member of the Brucella family as well as providing a marker for select *Brucella* linages. The interconnection of sulfur metabolism and the capacity for *Brucella ovis* to survive its preferred environment is interesting to consider in the search for novel targets for antibiotics. It is also another window into the pathways that are involved in *Brucella* adaptation to hostile environments where it has evolved to thrive. Dissecting the nature of the response of *Brucella* to the intracellular niche, specifically the different stressors and growth phases, are increasing our understanding on how bacteria adjust to change and combat specific attacks.

My work has provided insight into the reactions to the different types of environments *Brucella ovis* encounters during its life cycle. As a facultative intracellular bacterium, it is thoroughly challenged by the host cell environment and has thus evolved mechanisms that permit

it to thrive even in this hostile habitat. Sensing, persisting and reacting to its everchanging surroundings have enabled this furtive bacterium to establish a niche that ensures its survival and proliferation, granting it all the tools it requires to endure and evolve with, and perhaps beyond, its host.

5. APPENDIX A – REGULATION OF THE *ERYTHROBACTER LITORALIS* DSM 8509 GENERAL STRESS RESPONSE BY VISIBLE LIGHT

5.1. PREFACE

The contents of this appendix was adapted from its published form:

Fiebig A, **Varesio LM**, Alejandro Navarreto X and Crosson S. *Molecular Microbiology* (2019).

 Molecular Microbiology 112(2), 442–460 DOI:10.1111/mmi.14310

Additional supplementary information (**Tables S1-S6**) may be found online.

5.2. INTRODUCTION

5.2.1. LOV-HWE kinases: An overview

Light is a ubiquitous environmental factor that provides energy to support life in many ecosystems. Multiple adaptive responses to light have evolved in bacteria including phototaxis and photoavoidance, photoprotective pigment production, and the regulation of genes required for photosynthesis. These responses are mediated by proteins that sense photons in particular energy ranges across the visible spectrum (272). LOV domains are widely distributed photosensors (273) that detect blue light via a bound flavin cofactor (274). These photosensory domains are present in an assortment of signal transduction proteins from bacteria, archaea, fungi, protists, and plants (275, 276). Though broadly conserved (277, 278), the physiological roles of LOV photosensors in bacteria remain largely undefined.

The class *Alphaproteobacteria* contains species that inhabit diverse ecological niches, and that have a significant impact on nutrient cycling, agricultural production and human health (279). *Alphaproteobacteria* commonly encode proteins that contain a LOV domain coupled to an HWE/HisKA2-type (280) histidine kinase. Surprisingly, these "LOV-HWE kinases" are often present in heterotrophic species with no evident photobiology (276). Histidine kinases are typically co-expressed with and phosphorylate their cognate response regulators to directly control gene expression (281), but HWE/HisKA2 kinases are unusual in that they are often orphaned on the bacterial chromosome, or are adjacent to single domain response regulators (SDRR) that lack a regulatory output domain to control gene expression (280).

Early studies of Alphaproteobacterial LOV-HWE kinases in *Caulobacter crescentus* (282) and *Rhizobium leguminosarum* (283), demonstrated their influence on cell adhesion and biofilm formation. In the case of *C. crescentus*, a LOV-HWE kinase (LovK) modulates adhesion by controlling expression of a single downstream gene that regulates surface adhesin biosynthesis (284, 285). However, the major effect of deleting or overexpressing genes encoding LOV-HWE kinases appears to be dysregulation of the general stress response (GSR) system (246, 286), which determines cell survival across a range of stress conditions (40, 42). There is increasing evidence that multiple HWE/HisKA2-family kinases function as part of complex regulatory networks that regulate the GSR in *Alphaproteobacteria* by influencing the phosphorylation state of the anti-anti-σ factor, PhyR (287–290). Phospho-PhyR activates an extracytoplasmic function (ECF) σ factor – EcfG – by binding and sequestering its anti-σ factor, NepR (291) (**Figure 6.32**).

Though there is genetic evidence that LOV-HWE kinases control transcription of genes in the GSR regulon (246, 286, 287, 292), the effects of visible light on LOV-HWE kinase signaling *in vivo* remain undefined. Therefore, it is important to identify appropriate experimental systems

to assess the effect of visible light and other environmental signals on LOV-HWE kinase signaling in living cells.

5.2.2. *Erythrobacter litoralis* DSM 8509 as a system to study LOV-HWE kinase signaling

We sought to investigate the signaling role of LOV-HWE kinases in the genus *Erythrobacter*, which was attractive to us for several reasons. *Erythrobacter* spp. are abundant in the world's oceans (278, 293, 294) where they contribute to global nutrient and energy cycling (295, 296). Importantly, LOV-HWE kinases are common in *Erythrobacter* spp., as evidenced in both sequenced isolates (297, 298) and in a set of metagenome assembled genomes (MAGs) (299) from the Tara oceans sequence datasets (278). Though *Erythrobacter* spp. can be isolated and grown in axenic culture (298, 300), strains have not been extensively cultivated and manipulated. Thus, *Erythrobacter* spp. are less likely to have acquired lab-adaptive mutations that could potentially skew the effects of light on cell physiology. Most *Erythrobacter* spp. are aerobic anoxygenic photoheterotrophs (AAP), which carry a photosynthesis gene cluster that encodes production of bacteriochlorophyll *a* (Bchl *a*), and other genes required for phototrophy. Therefore, physiological responses to visible light are expected in this genus.

Erythrobacter litoralis strain DSM 8509 was isolated from a cyanobacterial mat in the supralitoral zone off the island of Texel, Netherlands (301, 302). Draft genome sequence of DSM 8509 indicated the presence of a single LOV-HWE kinase gene (298). This distinguishes DSM 8509 from *E. litoralis* strain HTCC 2594, which encodes multiple LOV-kinases (297) that have been previously characterized *in vitro* (290, 303, 304). HTCC 2594 does not encode the genes required for phototrophy (297, 305), and though DSM 8509 and HTCC 2594 cluster phylogenetically based on their 16S sequences, they do not group based on amino acid sequence

of concatenated core genes (305). As such, it has been suggested that HTCC 2594 should be re-classified as a distinct species, *Erythrobacter* sp. HTCC 2594 (306).

Here, we report the development of *E. litoralis* DSM 8509 as an experimental genetic system, which we have used to study the contribution of a LOV-HWE kinase to regulation of the general stress response. Our data show that an ensemble of three HWE kinases, including the LOV-HWE kinase LovK, collectively regulates the general stress response in *E. litoralis*. The cytoplasmic kinase, GsrK, functions as an activator of GSR transcription in complex medium, while the transmembrane kinase, GsrP, is a repressor. LovK activates GSR transcription, and is a more potent GSR activator under dark conditions. Transcription of the entire GSR regulon is significantly higher in dark conditions than in light. In fact, the set of genes that are differentially transcribed upon shifts in the light environment strongly overlaps the experimentally-defined GSR regulon. While LovK contributes to light-dependent regulation of GSR transcription, it is not strictly required for this response. Our results support a model in which photons directly and indirectly modulate GSR transcription via the HWE kinases GsrK, GsrP, and LovK.

5.3. RESULTS

5.3.1. A brief comparison of *Erythrobacter litoralis* DSM 8509 to other *Erythrobacter* species

Several partial and complete genome sequences of *Erythrobacter* isolates have been deposited in public databases (see Table 7.7 for representative genomes). These genomes have similar characteristics in terms of size and GC content, yet differ in presence/absence of particular genes and pathways. For example, *Erythrobacter* spp. can encode zero, one, or multiple LOV-HWE kinases, and the total number of HWE/HisKA2-family kinases ranges from two to sixteen (307). Although the three LOV kinases present in strain HTCC 2594 have been biochemically

characterized (290, 303, 304), we anticipated that the potential functional redundancy of these genes could complicate interpretation of genetic and physiological studies *in vivo*. In contrast, DSM 8509 encodes only one LOV-HWE kinase. Additionally, when considering our intent to investigate the interplay between multiple HWE/HisKA2-family kinases involved in GSR regulation, we noted that DSM 8509 has a 'goldilocks' number of such kinases: more than one, but not too many to complicate genetic analyses (Table 7.9). Finally, DSM 8509 encodes core genes required for phototrophy, which provides an opportunity to develop genetic tools in a species for which light is likely a central environmental/metabolic signal. For these reasons, we chose to pursue the development of *E. litoralis* DSM 8509 as a comparative model to study the function of LOV-HWE kinases.

5.3.2. Whole-genome sequencing and development of genetic tools in DSM 8509

To our knowledge, genetic analysis of an *Erythrobacter* species has not been reported. To support development of *Erythrobacter litoralis* DSM 8509 as an experimental genetic model, we first completed and closed the genome sequence. Specifically, we re-sequenced DSM 8509 using a PacBio long-read sequencing platform to an average depth of 143X (**Table S1**). These sequence reads were assembled *de novo* to a single 3.25 Mb contig, which was deposited and is available to the public under NCBI GenBank accession CP017057. We surveyed the antibiotic sensitivity profile of DSM 8509 to identify useful markers for genetic selection (**Table S2**), and adapted methods to transform the bacterium with both replicating (pBBR-derived) and integrating (ColE1-derived) plasmids. We successfully transformed *E. litoralis* DSM 8509 by both electroporation and conjugation (see **Materials and methods**). We further identified conditions to generate unmarked gene deletion and allele replacement strains using a two-step recombination and

sacB/sucrose counterselection approach, which we applied to our genetic analysis of HWE kinases and GSR signaling described below.

5.3.3. *E. litoralis* DSM 8509 LovK is a photosensor

Like *Caulobacter crescentus* (282, 308), *E. litoralis* DSM 8509 possesses a single LOV-HWE histidine kinase gene that is located adjacent to a single-domain response regulator gene on the chromosome (locus tags Ga0102493_111685 and 111686). We have named these genes *lovK* and *lovR*, respectively (**Figure 6.33**). The N-terminal LOV domain of LovK contains a complete flavin binding consensus sequence (276, 309) and a conserved cysteine residue required for light-dependent cysteinyl-flavin adduct formation (310). To validate that LovK indeed functions as a *bona fide* photoreceptor *in vitro*, we cloned and expressed LovK in a heterologous *E. coli* system, and purified the protein by affinity chromatography. LovK has a classic LOV domain visible absorption spectrum, with a λ_{max} at 450 nm and vibronic bands at 425 nm and 475 nm. Illumination of the protein with blue light results in loss of these major bands, and a concomitant increase in absorption at 396 nm (**Figure 6.34**). This spectral signature is consistent with formation of a covalent cysteinyl-flavin C4(a) adduct upon illumination (310–313). We further expressed and purified LovK(C73A), in which the conserved cysteine postulated to form a flavin C4(a) adduct is mutated to a non-reactive alanine. LovK(C73A) retains flavin binding as evidenced by its absorption spectrum, but does not undergo the light-dependent bleaching of visible absorption bands that are indicative of cysteinyl-flavin adduct formation (**Figure 6.34**). These data provide evidence that *E. litoralis* DSM 8509 LovK is a photosensor that has the photochemical features of typical LOV proteins.

5.3.4. The *E. litoralis* DSM 8509 GSR regulon

Before testing whether LovK plays a role in the regulation of general stress response (GSR) transcription, we first sought to define the GSR regulon in *E. litoralis* DSM 8509 by RNA-sequencing (RNA-seq). To this end, we deleted the predicted orthologs of *phyR* (locus tag Ga0102493_111538) and the *nepR-ecfG* operon (locus tags Ga0102493_111536 and 111535) (**Figure 6.33**). As discussed in the introduction, these genes encode a protein partner switching system that controls activation of GSR transcription by σ^{EcfG}. Deletion of either *phyR* or *ecfG* were predicted to result in a strain incapable of activating GSR transcription. We compared the global transcriptional profiles of the Δ*phyR* and Δ*nepR-ecfG* deletion strains to wild type (WT) (**Table S3**). To examine the effects of light, each strain was grown in cool white fluorescent light (60 μmol m^{-2} s^{-1}) or in darkness (foil-covered tubes) for 24 hours prior to harvesting mRNA.

A shared set of 183 genes were differentially expressed (more than 1.5-fold; false discovery rate (FDR) p-value < 0.01) in Δ*phyR* and Δ*nepR-ecfG* compared to wild-type cells (**Figure 6.35**, Table S4). With few exceptions, transcript levels in this gene set were lower in the two mutant strains compared to wild type indicating that these differentially regulated genes are transcriptionally activated by σ^{EcfG} in wild-type cells. We identified a predicted σ^{EcfG} binding motif that matched EcfG motifs from other *Alphaproteobacteria* (40–42, 314) in the promoter region of approximately one-quarter of the regulated genes (**Figure 6.35**, **Table S4**). Presence of these motifs predict the subset of genes in the *E. litoralis* GSR regulon that are directly regulated by the ECF sigma factor, σ^{EcfG}.

Genes encoding transporters, cell envelope, and cell surface proteins are broadly represented in the *E. litoralis* DSM 8509 GSR regulon (**Table S4**). Transcription of genes encoding the ferritin-like protein DPS, a SOUL heme-binding protein, and superoxide dismutase (*sod*) are commonly under the control of the GSR system in other *Alphaproteobacteria* (39, 222,

246, 286, 291, 315–321), and are regulated by *nepR-ecfG* and *phyR* in *E. litoralis* DSM8509 (**Table S4**). The promoters of *dps, sodA, sodC*, and a predicted SOUL heme-binding protein gene (Ga0102493_112016) contain EcfG motifs and thus appear to be directly activated by σ^{EcfG}. Our transcriptomic data also provide evidence that expression of several DNA photolyases is activated by the GSR system.

Genes directly or indirectly regulated by σ^{EcfG} that may be relevant to the photobiology of *E. litoralis* include a gene encoding a TspO/MBR tryptophan-rich protein (locus Ga0102493_112559). This class of outer membrane protein has been reported to bind tetrapyrroles and promote their photooxidative degradation (322), and to negatively regulate photosynthesis in *Rhodobacter* (323). Transcripts levels from this particular TspO/MBR gene are highly reduced in strains lacking *phyR* or *nepR-ecfG*. Unlike the fungus *Aspergillus fumigatus*, where TspO/MBR expression is strongly induced by light exposure (324), expression of *E. litoralis* TspO/MBR (Ga0102493_112559) is lower in light than in dark conditions. Notably, steady-state transcript levels from the bacteriochlorophyll biosynthesis gene, *bchC*, are higher in both Δ*phyR* and in Δ*nepR-ecfG* relative to wild type (**Table S4**). Transcripts of other bacteriochlorophyll biosynthesis genes including *bchF*, *bchB* and *bchN* are also higher in Δ*phyR* and Δ*nepR-ecfG* strains, though to a lesser extent than *bchC* (**Figure 6.36**). These results indicate that the *E. litoralis* GSR system directly or indirectly represses select genes involved in phototrophy.

5.3.5. Light-dependent regulation of transcription at the genome scale

In wild-type *E. litoralis*, transcripts in the GSR regulon are more abundant in dark-grown cells (aluminum foil-covered tubes) than in cells illuminated with white light (60 μmol m^{-2} s^{-1}) (**Figure 6.35**). In fact, measured GSR transcript levels from illuminated wild-type cells are only slightly higher than mutant strains lacking either *phyR* or *nepR-ecfG*. From this result, we conclude

that the GSR is only marginally active when cells are grown in continuous white light (**Figure 6.35 and Table S4**). A light versus dark difference in GSR transcripts is not observed in Δ*phyR* or Δ*nepR-ecfG* cells.

In wild-type cells, the majority of genes differentially expressed between light and dark conditions are in the GSR regulon. Of the 220 transcripts that change more than 1.4-fold (FDR p-value < 0.01), only 37 are not in the GSR regulon as defined (**Figure 6.37, Table S5**). Twenty-five of the differentially-expressed genes outside of the GSR regulon have similar light-dependent changes in the Δ*phyR* and Δ*nepR-ecfG* strains, and thus represent a GSR-independent light-regulated gene set. This GSR-independent gene set exhibits modest (less than 2-fold) regulation in response to white light treatment, and includes photosynthesis genes with reduced transcript levels in the light (**Figure 6.36 and Figure 6.38**). These genes are distinct from those derepressed in the GSR mutant strains (e.g. *bchC*) and include *bchH*, *bchL*, *puhA*, and *bchM*. We conclude that there are genes involved in photosynthesis (e.g. *bch*, *puf*, *puh*) for which expression is influenced by the GSR system, and others for which regulation is independent of GSR (**Figure 6.36**).

Reduction in photosynthesis gene expression in *E. litoralis* in the light is consistent with light repression of reaction center genes in the purple photosynthetic bacterium, *Rhodobacter capsulatus* (325), and downregulation of photosynthesis-related genes in the aerobic anoxygenic phototrophic bacterium, *Dinoroseobacter shibae*, upon a shift from heterotrophic growth in the dark to photoheterotrophic growth in the light (326). An operon of unknown function (Ga0102493_112844-50) involved in acyl-CoA metabolism is repressed by light, independent of the GSR system. Eight metabolic genes of varying function are activated by light, independent of the GSR system (**Figure 6.37**).

5.3.6. Regulators of the GSR signaling pathway

Typically, transcription of genes encoding the core GSR regulators NepR, EcfG, PhyR and GSR sensory kinases is activated by σ^{ecfG} (222, 246, 286, 291, 315, 316, 319, 320). In *E. litoralis*, a σ^{EcfG}-binding motif is positioned directly upstream of the *nepR-ecfG* operon. As expected, *nepR-ecfG* transcripts were reduced in the strain lacking *phyR* and in cells grown in the light (**Figure 6.33C**, **Table S4**).

It was previously noted that conserved nucleotides at -35 and -10 within the σ^{EcfG}-binding motif are nearly palindromic, which lead to the hypothesis that a palindromic site might function to drive bi-directional expression of oppositely oriented genes (314). The -35 and -10 sites upstream of *nepR-ecfG* have a strong palindromic character (GGAAC-N_{17}-GTTCC) (**Figure 6.38, Table S4**). Our transcriptomic data provide evidence that expression of both *nepR-ecfG* and the oppositely oriented HWE sensor kinase gene (Ga0102493_111537, *gsrP*) is activated from this single palindromic site. Both *nepR-ecfG* and *gsrP* are part of the GSR regulon as determined by RNA-seq (**Figure 6.33C, Table S4**). Moreover, RNA-seq reads corresponding to *nepR-ecfG* and *gsrP* transcripts each begin 13-14 bp downstream of either end of this shared promoter motif (**Figure 6.38**). While this is not a definitive mapping of transcriptional start sites, these data strongly suggest that this particular motif can functional bi-directionally. The mechanism by which σ^{EcfG} could regulate transcription from both strands at this site is not known. We note that the motif oriented toward *nepR-ecfG* has a stronger score in MEME, and that more RNA-seq reads mapped to *nepR-ecfG* than *gsrP*, which together suggest that σ^{EcfG} preferentially initiates transcription of *nepR-ecfG*. Finally, we observed a small number of RNA-seq reads that mapped to the negative strand of this motif and its downstream region (genome positions 619,592-619,621, **Figure 6.38**). We speculate that these reads may correspond to transcription initiation from a cryptic or non-ECF

σ promoter that would ensure some minimal level of *nepR-ecfG* expression under conditions in which the GSR was completely inactive.

In contrast to many previously described systems, an EcfG motif was not evident in the *phyR* promoter and, accordingly, transcript levels for this gene were not affected by deletion of *nepR-ecfG* or by light conditions (**Figure 6.33C and Figure 6.38, Table S4**). In other words, *phyR* appears to be constitutively expressed in *E. litoralis* DSM 8509. Constitutive expression of *phyR* likely necessitates the presence of a negative PhyR regulator to avoid constitutive activation of the GSR.

HWE/HisKA2-family kinases are associated with the GSR regulatory system in *Alphaproteobacteria* (40, 42, 327). *E. litoralis* DSM 8509 encodes five HWE/HisKA2-family kinases. Of these, only three are expressed at an appreciable level under our laboratory cultivation conditions, including: *1)* the kinase encoded adjacent to *nepR-ecfG* operon (Ga0102493_111537), *2) lovK*, and *3)* an orphan sensor kinase (Ga0102493_11718) (**Figure 6.33**). Ga0102493_111537 encodes a transmembrane HWE sensor kinase with a periplasmic CHASE sensory domain; its transcription is activated by *phyR* and *ecfG* and by dark conditions (**Figure 6.33C and Figure 6.38**). As noted above, a shared EcfG motif identified between *nepR-ecfG* and this kinase likely promotes expression of both transcripts. The *lovK-lovR* promoter also contains an EcfG motif, and transcription of *lovK-lovR* requires *nepR-ecfG* and *phyR* (**Figure 6.33C**). The orphan HWE kinase, Ga0102493_11718, lacks an EcfG motif in its promoter and is constitutively expressed in our experimental conditions (**Figure 6.33C**). Of the remaining HWE-kinase genes, transcripts corresponding to the orphan kinase gene Ga0102493_112963 are nearly undetectable (RPKM < 10). HWE-kinase gene, Ga0102493_111751, is encoded in an operon with a CRP-family

transcription factor and a single domain response regulator. Transcripts for these genes are also low abundance (**Figure 6.39**).

5.3.7. The role of three HWE-family sensor kinases in regulation of GSR transcription

To assess the role of HWE-family kinases 111537, 11718, and LovK in regulation of GSR transcription, we constructed strains bearing in-frame unmarked deletions of each kinase gene. We then evaluated GSR transcriptional output in these strains by measuring levels of a GSR-regulated transcript, *dps*, by qRT-PCR. This transcript was selected because *1)* it exhibits a large dynamic range of expression between "GSR-ON" and "GSR-OFF" conditions (**Figure 6.40**), *2)* it contains an EcfG motif in its promoter suggesting direct regulation by σ^{EcfG} (**Table S4**), and *3)* primers to this transcript provided efficient amplification in one-step qRT-PCR reactions (see **Materials and methods**). Gene locus Ga0102493_112759, which encodes a methylmalonyl-CoA mutase, was selected as a control gene for normalization (**Figure 6.40**). In our RNA-seq data sets, transcripts for this gene were abundant (in the top 20% of all transcripts) and the RPKM coefficient of variation between samples was among the lowest, making this a suitable control gene for normalization. We tested these qRT-PCR primer sets in the same strains evaluated by RNA-seq and patterns of *dps* expression were consistent between these two methods (**Figure 6.40**).

In a strain lacking kinase 111537, steady-state levels of *dps* transcript increased in cultures grown both in the light and in the dark (**Figure 6.41A**). These data are consistent with 111537 functioning as a negative regulator of GSR transcription. Conversely, deletion of the 11718 resulted in decreased *dps* transcript levels, consistent with this gene functioning as a positive regulator of GSR transcription. Both of these transcriptional phenotypes were complemented by expression of the deleted gene from its native promoter on a replicating plasmid (**Figure 6.42**). We have named these genes *gsrP* and *gsrK*, for *g*eneral *s*tress *r*esponse *p*hosphatase and *k*inase

respectively. We note that the putative phosphatase and/or kinase activity of these proteins has not been established, but our genetic data are consistent with these biochemical activities against PhyR.

Deletion of the *lovKR* operon did not have a significant effect on *dps* transcript levels as measured by qRT-PCR (**Figure 6.41A**). Moreover, we did not observe large differences in GSR transcription in a Δ*lovKR* strain at the genome scale compared to wild type in our RNA-seq experiments (**Figure 6.35**). Nonetheless, GSR transcripts were uniformly lower in Δ*lovKR* compared to wild type when strains were grown in the dark (**Figure 6.35D and E**). The average $\log_2$ (fold change) difference in all GSR transcripts between these strains is -0.25, which reflects ~15% reduction in GSR transcription in the Δ*lovKR* strain ($p < 0.0001$). This trend is specific to the GSR regulon: the average $\log_2$(fold change) difference for all transcripts is 0.01 reflecting that, on average, the transcriptome at the whole-genome scale remains unchanged. These results provide evidence that *lovKR* plays a subtle role as an activator of GSR transcription in the dark.

To examine possible redundancy in the functions of these HWE kinases, we evaluated the effect of deleting pairs of kinases, leaving one of the three genes intact. In a Δ*gsrK* Δ*lovKR* double deletion strain where *gsrP* remains, *dps* transcript levels are lower than either of the single deletion strains in both the light and dark (**Figure 6.41A**). This result supports a model in which both GsrK and LovK function as GSR activators; moreover GsrP does not activate GSR transcription in the dark or in the light. In the Δ*gsrP* Δ*lovKR* double deletion strain where *gsrK* remains, *dps* transcripts are elevated in both the light and the dark, similar to the strain lacking only *gsrP*. Finally, in the Δ*gsrP* Δ*gsrK* strain where *lovKR* remains, *dps* transcripts are higher than wild type in both the light and the dark (**Figure 6.41A**) and are differentially expressed in response to light treatment. In this strain, *dps* levels are higher in dark-grown than light-grown cells. To further investigate

whether LovK functions as a global activator of GSR, we measured the transcriptome of the Δ*gsrP* Δ*gsrK* strain grown in the light and in the dark by RNA-seq. These experiments confirm the single gene (*dps*) results in this strain. Specifically, GSR transcripts in this strain are broadly elevated compared to wild type in either light condition, and are higher in the dark than the light (**Figure 6.35, Table S4**). Together these data support a model in which LovK can function as an activator of GSR transcription that has enhanced activity in the dark.

To test whether LOV domain photochemistry (i.e. cysteinyl-flavin covalent adduct formation) contributes to the light-dark difference in LovK-dependent transcription, we mutated the conserved LOV domain cysteine (C73) to an alanine in the Δ*gsrP* Δ*gsrK* strain. Purified LovK(C73A) mutant protein is blind to light (**Figure 6.34**), and in the Δ*gsrP* Δ*gsrK lovK(C73A)* strain, GSR transcription was no longer sensitive to light (**Figure 6.41B**). Specifically, *dps* levels were comparably high in light and dark grown cultures and similar to dark grown Δ*gsrP* Δ*gsrK* *lovKR*$^{+}$ cultures. From these data, we again conclude that LovK is a more potent activator of GSR in its dark state.

We further tested whether the conserved histidine phosphorylation site in LovK (H161) was required to regulate GSR transcription. *dps* transcript levels in the Δ*gsrP* Δ*gsrK lovK(H161A)* strain are similar to a strain lacking all three kinases (Δ*gsrP* Δ*gsrK* Δ*lovKR)* indicating that a *lovK(H161A)* mutant behaves like a Δ*lovKR* deletion. We conclude that LovK phosphorylation is necessary for LovK to activate GSR transcription (**Figure 6.41B**). Together, the data provide evidence that *gsrP*, *gsrK* and *lovK* coordinately regulate GSR transcription under our assayed conditions.

5.4. DISCUSSION

Regulation of transcription by alternative sigma factors including σ^S in *Gammaproteobacteria* (11, 328), σ^B in select Gram-positive bacteria (35), and σ^{EcfG} in *Alphaproteobacteria* (40, 42) confers general resistance to a range of physical and chemical stress conditions *in vitro*. Though these general stress response (GSR) σ factors are conserved within phylogenetic groups, the input signals that cue their activation and the σ-dependent transcriptional outputs vary across species. GSR inputs and outputs presumably reflect the distinct physicochemical challenges that particular species encounter within their niches. In this study, we report the development of the aerobic anoxygenic photoheterotroph (AAP), *Erythrobacter litoralis* DSM 8509, as a comparative genetic model system to study regulation of GSR transcription by σ^{EcfG}. More specifically, we have sought to define the regulatory role of HWE-family sensor histidine kinases (280)— including a photoresponsive LOV-HWE kinase — in σ^{EcfG}-dependent transcription in *E. litoralis*.

5.4.1. Light, LOV, and bacterial stress responses

Proteins containing photosensory LOV domains are widely distributed in archaea, eukarya, and bacteria (275). The possibility that visible light regulates bacterial stress responses via LOV domains was first noted (273)after the discovery of the *Bacillus subtilis* LOV-STAS protein, YtvA, which functions as an activator of σ^B-dependent transcription (329). It was later shown that blue light can indeed activate σ^B-dependent transcription in *B. subtilis* via YtvA (330). Subsequent studies of a related LOV-STAS protein from *Listeria monocytogenes* support the conclusion that light regulation of σ^B by LOV proteins occurs more broadly in Gram-positive bacteria (331, 332). A regulatory link between LOV and σ^S has also been reported in *Pseudomonas syringae*, where

white light was shown to repress expression of the gene encoding σ^S (*rpoS*) via a LOV histidine kinase (333).

Though the Alphaproteobacterial GSR sigma factor, σ^{EcfG}, is not related to σ^B or σ^S, studies of *Caulobacter crescentus* LovK provide evidence that LOV-HWE kinases repress σ^{EcfG}-dependent transcription. Experiments in *Brucella abortus* have identified a related LOV-HWE kinase, LovhK, that activates σ^{EcfG}-dependent transcription (286). However, these studies have not provided evidence that visible light is an input signal that influences σ^{EcfG}-dependent transcription (246, 292). In *E. litoralis* DSM 8509, we have shown that the activity LovK as a regulator of σ^{EcfG} is influenced by light. More specifically, LovK functions with two additional HWE-family kinases, GsrP and GsrK, to control GSR output (**Figure 6.43A**): GsrK is a strong GSR activator, GsrP is a repressor, and LovK is an activator that can be directly modulated by light (**Figure 6.43A**). These data contribute to an emerging model of GSR regulation in *Alphaproteobacteria* in which consortia of HWE/HisKA2-family kinases regulate σ^{EcfG}.

Our RNA-seq experiments show that *E. litoralis* GSR transcription is highly activated in the dark, and only marginally activated in cells continuously illuminated with white fluorescent light. In contrast, light transiently induces transcription of the core GSR regulators (Dshi_3834-3837) and other stress response genes in the related AAP species, *Dinoroseobacter shibae* (326). Similarly, blue light activates *rpoE*- and *rpoH*-dependent stress responses in the phototrophic Alphaproteobacterium *Rhodobacter sphaeroides* (334, 335). We did not measure transient responses of *E. litoralis* to light or dark shifts, and the advantage (if any) of elevated GSR transcription in the dark is not known.

5.4.2. LovK functions as part of a consortium of GSR sensor histidine kinases

Although LovK can clearly function as a photoreceptor, it is not explicitly required for light/dark control of GSR transcription. Specifically, transcription from a GSR-regulated promoter remains light-responsive in a *gsrK*$^{+}$ *gsrP*$^{+}$ Δ*lovKR* strain. Thus, the combined activities of GsrK and GsrP must be balanced in a manner that leads to higher transcriptional output in dark relative to illuminated conditions (**Figure 6.41A**). This light responsiveness could emerge from enhanced activity of GsrK in the dark, enhanced repression by GsrP in the light, or a combination of both. There is no evidence to suggest that photons are a direct signal for either sensor, thus we predict that either GsrK and/or GsrP indirectly sense a metabolic or physiological change that occurs upon shifts in the light environment.

We favor a model in which GsrP is a more potent GSR repressor in cells grown in the light (**Figure 6.43A**). In our experimental conditions, *gsrK* alone strongly activates GSR transcription in light and dark (**Figure 6.41A**). This gene is constitutively expressed in both conditions (**Figure 6.33C**). While our results suggest that GsrK activity is not affected by light, we cannot entirely exclude the possibility that light conditions influence GsrK. GSR output from the strain bearing only *gsrP* is equivalently low in light and dark conditions (**Figure 6.41A**). While light may influence the activity of GsrP as a repressor, it is difficult to assess this possibility in the absence of a kinase that activates GSR transcription. σ^{EcfG}–dependent transcription of *gsrP* comprises a negative feedback loop that controls GSR transcriptional output; the level of *gsrP* transcripts is dramatically reduced in light conditions (**Figure 6.33C, Table S4**). This regulatory circuitry leads one to predict that the activity of GsrP as a repressor is inversely proportional to its steady-state levels in the cell. The activities of *E. litoralis* LovK, GsrK, and GsrP likely serve to both counteract and reinforce each other depending on environmental conditions. The possibility that *E. litoralis* and other *Alphaproteobacteria* can coordinately perceive multiple environmental inputs through a

consortium of HWE-family kinases may endow these species with the ability to execute complex decision-making processes with regard to control of GSR transcription.

Finally, we note that our data support a model in which LovK is active in both the light and dark, though activity of LovK as a positive regulator of σ^{EcfG}-dependent transcription is enhanced in the dark. These *in vivo* data are consistent with *in vitro* data showing that light does not regulate *E. litoralis* HTCC 2594 LOV kinases in a binary (i.e. ON-OFF) manner, but rather modulates phosphorylation kinetics (290, 336). Though blue light typically enhances the enzymatic activity of LOV kinases *in vitro*, there is precedent for the unlit, dark state being the active state *in vivo*. In *D. shibae,* a short LOV protein is a dark activator of pigment biosynthesis (337). In this case, the LOV domain may be thought of as a dark sensor rather than a light sensor.

One may speculate on reasons why LOV proteins are often incorporated into GSR systems in the bacterial kingdom, but the fact remains that we have little understanding of the physiological relevance of LOV-regulated responses to changes in the light or redox environment in bacteria. *E. litoralis* DSM 8509 is a tractable experimental system with interesting physiological features. This species can now be leveraged as a comparative model to investigate core molecular mechanisms underlying environmental stress responses in *Alphaproteobacteria*.

5.5. MATERIALS AND METHODS

Growth of E. litoralis

Erythrobacter litoralis DSM 8509 was obtained from the American Type Culture Collection (ATCC 700002). For these studies, this organism was grown in Difco Marine Broth 2216. When broth was prepared according to manufacturer's instructions, cells tended to clump in flocs. Dilution of marine broth to 0.5X reduced flocculation, and thus all cultures were grown at this

broth dilution (18.7 g of powder per L). 0.5X marine broth was sterilized by autoclaving and, prior to inoculation with *E. litoralis*, the medium was passed through a sterile 0.22 μm filter to remove precipitates. Liquid cultures were grown in an Infors air incubator at 30 ˚C in glass culture tubes, inclined at a 45˚ angle, shaking at 200 RPM. For "light" conditions, a bay of 10 fluorescent Philips and Sylvania T8 lights inside the shaking incubator was switched on. For "dark" conditions glass tubes were carefully wrapped in aluminum foil. For molecular genetic manipulation and isolation of mutants, colonies were grown on 0.5X marine broth solidified with 15 g of agar per L of medium. Agar plates were grown in a 30 ˚C air incubator, or at room temperature in ambient light.

Growth of Escherichia coli *strains*

E. coli strains were grown in LB Miller medium (10 g peptone, 5 g yeast extract, 10 g NaCl per L) at 37˚C. Growth medium was solidified with 1.5% agar. Antibiotics were added at the following concentrations as appropriate: kanamycin, 50 μg/ml; gentamycin 15 μg/ml; chloramphenicol 12 μg/ml.

Antibiotic sensitivity testing

To evaluate the appropriate antibiotic concentrations to use for selection in genetic manipulations, *E. litoralis* DSM 8509 cells were first grown in a streak on 0.5X marine agar at 30 ˚C for 2-3 days. Cells were scraped from the agar plate into 1 ml of 0.5X marine broth and evenly suspended by pipetting up and down. The optical density (OD) at 660 nm was was adjusted to ~ 0.1 AU. Cells were then 10-fold serially diluted and 20 μl of each dilution was spotted onto 0.5X marine broth agar that had been supplemented with a range of antibiotics concentrations. The antibiotic concentrations tested were as follows: 200, 100, 50, 25, 10 μg/ml for ampicillin, carbenicillin and tetracycline; 100, 50, 20, 10, 5, 1 μg/ml for gentamycin, spectinomycin, streptomycin, apramycin,

and rifampicin. After the liquid absorbed into the agar, the plates were incubated at 30 ˚C for 1 week. The minimal antibiotic concentrations that resulted in at least 4 orders of growth inhibition are reported in **Table S2**.

Molecular cloning for plasmid generation

Plasmids for expression or allele replacement were generated using routine techniques. Genes or loci of interest were PCR amplified with KOD Xtreme Hot Start polymerase (Millipore Sigma) using the primers listed in Table S6. To aid amplification of these GC-rich sequences, PCR reactions were supplemented with 5% DMSO (final concentration). Primers included a 5' extension with restriction endonuclease sites or overhangs for overlap extension PCR reactions. Null alleles were generated by first amplifying fragments ~500 bp upstream and ~500 bp downstream of the gene of interest. These two fragments were "stitched" together with overlap extension PCR to generate the desired allele. PCR products were cleaned using GeneJet DNA cleanup columns (Thermo Fisher), digested with appropriate restriction endonucleases (NEB) to generate overhang sequences, and ligated into similarly digested plasmids using T4 DNA ligase (NEB). Ligation reactions were transformed into chemically competent TOP10 *E. coli* by heat shock. Transformants were selected on LB supplemented with the appropriate antibiotic and grown overnight at 37 ˚C. The inserted sequence and cloning junctions of all plasmids were confirmed by PCR amplification with plasmid specific primer sequences, treatment with ExoSapIT (Applied Biosystems) followed by Sanger sequencing (University of Chicago Comprehensive Cancer DNA Sequencing and Genotyping Facility). Plasmids generated for and used in this study are listed in **Table S6**.

Transformation of E. litoralis

We evaluated and optimized two methods for introducing DNA into *E. litoralis* DSM8509, electroporation and conjugation. Replicating plasmids (with BBR or RK2 replication origins) could be introduced by electroporation, but this method was not efficient enough to introduce non-replicating (i.e. suicide) plasmids for chromosomal integration. Conjugation, by triparental mating, was used to introduce suicide plasmids containing *oriT* sequences, and also replicating plasmids containing *oriT* sequences. Gentamycin (10 μg ml^{-1}) was used to select for clones carrying pBVMCS-4 derived plasmids. Chloramphenicol (1 μg ml^{-1}) was used to select for clones carrying integrated plasmids, which permitted us to generate in-frame deletion or allele replacement strains. Detailed electroporation and conjugation protocols follow.

Electroporation: Cells were scraped from a freshly grown plate that had been grown for 3 days at 30 ˚C. A pellet of 150-200 μl of cells was suspended in 750 μl 0.5X marine broth. Cells were pelleted by centrifugation (1 min at 14,000 x g) at 4 ˚C. The cell pellet was washed 3 times with 750 μl ice cold sterile water with a 1 min centrifugation step at each wash. After the final wash, the cells were resuspended with 140 μl cold water. The total volume of cells and water was ~200 μl. 60 μl cells were mixed with 0.1-1 μg purified plasmid and placed in a 1 mm electroporation cuvette. Cells were subjected to a single 1.8 kV pulse using the EC1 setting on a MicroPulser (Bio-Rad). Time constants were ~ 4-5 msec. 450 μl of 0.5X marine broth was added to the cuvette and cells were transferred in this broth to a sterile 13 mm glass culture tube and incubated at 30 ˚C shaking at 200 rpm for 2-4 hours. After this outgrowth, cells were plated on selective medium (about 200 μl per 100 mm petri dish). These plates were incubated at 30 ˚C for about 1 week. After 5 days, pinprick-sized colonies were visible. After 7-8 days, colonies were picked and struck on fresh selective medium to confirm antibiotic resistance and to select away any non-transformed

cells. We confirmed that the clones carried the plasmid of interest by colony PCR, amplifying plasmid specific sequences using cells as the template.

Conjugation: Plasmids encoding *oriT* sequences were transferred from *E. coli* TOP10 donor strains to E. *litoralis* using a helper strain (MT607 / pRK600) (338) which encodes the pilus required to mobilize *oriT* containing sequences. *E. litoralis* recipient strains were inoculated from fresh plates into 2 ml 0.5x marine broth in 13 X 100 mm glass culture tubes and grown shaking at 30˚C overnight. The *E. coli* TOP10 donor strain carrying the plasmid to be transferred was grown in 2 ml LB supplemented with the appropriate antibiotic and the *E. coli* helper strain was grown in 2 ml LB supplemented with chloramphenicol at 37˚C overnight. The three cultures were mixed at a ratio of 1 ml recipient, 200 μl donor, 200 μl helper. The mixed cells were centrifuged for 1 min at 10,000-14,000 x g. The pellet was resuspended in 50-100 μl 0.5x marine broth and then spotted onto a fresh 0.5x marine broth agar plate. After the liquid in the spot of cells absorbed into the agar, the plate was incubated overnight at room temperature or 30 ˚C. The mix of cells was scraped from this non-selective plate and spread on 0.5X marine broth agar supplemented with *1)* the appropriate antibiotic to select for acquisition of the plasmid and *2)* 100 μg /ml nalidixic acid (Nal) to counterselect against the *E. coli* donor and helper strains. *E. coli* strains exhibit increased resistance to Nal when grown on marine broth agar compared to LB, which required increasing the Nal concentration above typical concentrations used for counterselection. These selective plates were incubated for 7-8 days at 30 ˚C. Colonies that emerged were struck on fresh selective 0.5x marine broth agar plates and transformants were checked using the same approach described above for electroporated transformants.

Two-step chromosomal allele replacement

A standard two-step allele replacement approach with sucrose/*sacB* counterselection (339, 340) was optimized for *E. litoralis* DSM 8509. pNPTS138-derived allele replacement plasmids were introduced by conjugation as described above. Chloramphenicol resistant clones were inoculated into 2 ml 0.5X marine broth and allowed to grow without selection for 8-24 hours. These cells were spread on 0.5X marine broth agar supplemented with 7.5% sucrose to select for clones in which a second recombination event excised the plasmid. After approximately 1 week of growth at 30 ˚C, sucrose resistant colonies were replica patched on agar plates with or without chloramphenicol. From the chloramphenicol sensitive clones (i.e. clones in which the plasmid had excised) the locus of interest was PCR amplified and sequenced to distinguish clones bearing the parental allele from those bearing the desired new allele.

PacBio whole-genome sequencing

E. litoralis DSM 8509 genomic DNA was extracted using guanidium thiocyanate as previously described (341). Standard Pacific Biosciences (PacBio) large insert library preparation was performed. Briefly, DNA was fragmented to approximately 20kb using Covaris G tubes. Fragmented DNA was enzymatically repaired and ligated to a PacBio adapter to form the SMRTbell template. Templates larger than 10kb were BluePippin (Sage Science) size selected, annealed to sequencing primer, bound to polymerase (P6), bound to PacBio Mag-Beads and SMRTcell sequenced using C4 chemistry. The genome was sequenced using two SMRT cells and was assembled de novo using HGAP3 and polished with quiver. This process yielded a single, closed contig. The genome was automatically annotated using the DOE-JGI pipeline (342). HWE/HisKA_2-family histidine kinases and the core GSR regulators *phyR*, *ecfG*, and *nepR* were identified manually and annotated before sequence submission to GenBank. Raw PacBio reads

can be accessed through the NCBI sequence read archive at accession SRS1630618. The complete, annotated sequence of *E. litoralis* DSM 8509 is available at GenBank accession CP017057.

RNA extraction

For RNA-seq and qRT-PCR, RNA was extracted using a Trizol reagent based protocol. First, cells were struck from freezer stocks onto fresh 0.5X marine broth agar plates. After 3 days of growth, a scoop of mixed colonies of cells were inoculated into 2 ml 0.5X marine broth in 13 x 100 mm culture tubes and grown for 24 hours 30˚C shaking on a 45 degree incline at 200 RPM. Cultures were diluted to 0.005 OD_{660} in 2 ml fresh 0.5x marine broth and allowed to continue growth at 30 ˚C shaking at 200 RPM. After 20-22 hours, these cultures were diluted to 0.001 OD_{660} in 6 tubes (for RNA-seq) or 2 tubes (for qRT-PCR) with 2 ml each of 0.5X marine broth. Half of the tubes were placed in the top row of a shaker rack under a bay of 10 T8 fluorescent light bulbs (light conditions). The other half were wrapped in foil (dark conditions) and grown in the bottom row of the same shaker rack. After 22-24 hours of growth in constant light or constant dark conditions, culture densities reached 0.15-0.25 OD_{660}. Cells were harvested by rapid centrifugation in small "genotype-condition" batches to minimize handling time between culture growth and cell lysis. For RNA-seq, the 3 x 2 ml from each genotype-condition set were distributed over 4 x 1.5 ml microfuge tubes and centrifuged for 60-90 seconds at 14,000-17,000 x g. After quickly aspirating the supernatant liquid, the four pellets were resuspended in a total of 1 ml Trizol reagent (Invitrogen). For qRT-PCR, 1.5 ml of culture was harvested as above and resuspended in 1 ml Trizol. Cells lysed in Trizol were immediately stored at -80 ˚C until extraction. For RNA-seq, five samples of each genotype-condition were collected, each grown on a different day.

To extract RNA from the Trizol suspended cells, samples were thawed and incubated at 65˚C for 10 minutes. After addition of 200 µl of chloroform, samples were vortexed, incubated

on the benchtop for 5 minutes, then centrifuged for 12-15 minutes at 14,000-17,000 x g. The aqueous phase (~500 µl) was transferred to a new 1.5 ml tube and 450 ul isopropanol was added to precipitate the nucleic acid. Samples were frozen overnight at -80˚C then centrifuged at 17,000 x g for 30 minutes at 4 ˚C. The pellet was washed twice with 750 µl cold 70% ethanol, suspended in RNase free water (200 µl for RNA-seq samples and 100 µl for qRT-PCR samples), and stored at -80 ˚C.

RNA-seq

RNA was subjected to DNase digestion using TURBO DNase I (Thermo Fisher). RNA samples were bound to an RNAeasy column (Qiagen) and digested on the column by application of 70 µL of DNase coctail (7 µL DNase, 7 µL 10x Buffer, 56 µL diH_2O). Digests were incubated at room temperature for 45 minutes. RNA was cleaned and eluted from the column using buffers provided with the RNAeasy columns. DNA removal was confirmed by PCR using the RNA samples as a template and primers that amplify ~ 100 bp products (set 1: F-GACGGAGAAAAAGGCATCGC, R-GATTCGCCGTGTTCATCTGC; set 2: F-CCCACGAACCGATTTCATGG, R-CCTTCGGGGAGTTTCAAGCA). Absence of amplification confirmed removal of contaminating genomic DNA. Amplification from pre-DNAse samples served as a positive control. rRNA 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 manufacturer's instructions. The libraries were sequenced on an Illumina HiSeq 4000 by the University of Chicago Functional Genomics Core Facility. Data were analyzed using the RNA-seq workflow in CLC genomics workbench v11.0. RNA-sequencing reads have been deposited in the NCBI GEO database under accession GSE126532.

qRT-PCR

The GSR-dependent transcript, *dps* (Ga0102493_111653), and an endogenous control transcript (Ga0102493_112759) were evaluated using TaqMan probes and SuperScript III Platinum One-Step qRT-PCR Kit (Invitrogen) with a QuantStudio 5 real time PCR system (Thermo Fisher). The primer-probe sets (*dps* F – TCTCATCGCCGAACTCAAC; dps R – CGTGCCAGTGGAAATTCTTG; *dps* probe – /5HEX/CTTCGCGCT/ZEN/GTTCACCAAGACC/3IABkFQ/) and (ctrl F – AGATCGAAATGCTGTTGAAACG; ctrl R – GACCATCCAGAACGACATCC; ctrl probe -/56-FAM/CCGCAACAC/ZEN/CTATATCTACCCGCC/3IABkFQ/) were custom prepared by Integrated DNA Technologies. In control one-step RT-PCR reactions with a dilution series of template, both primer-probe sets exhibit 91-94% efficiency. Primers and probes were mixed to make a 10 μM F, 10 μM R and 5 μM probe stock solution. 0.4 μl primer-probe stock solution was used in each 20 μl reaction. 50 nM ROX served as the reference dye in each reaction. Each RNA sample was diluted to 2.5 ng μl^{-1}; 4 μl were used in each 20 μl reaction. Each sample was assayed in triplicate for each probe set. The average Ct from these technical replicates was considered the Ct for the sample. "No RT" control reactions were conducted on each sample with each probe to ensure that the signal from contaminating genomic DNA was less than 5 % of the signal (i.e. that the Ct from the "No-RT" reaction was at least 4 cycles later than the Ct from the "+RT" reaction. Two control samples (WT – Dark, and WT – Light samples) were diluted to 2.5 ng μl^{-1} and frozen in 50 μl aliquots. One aliquot was thawed and assayed on each plate to ensure consistency between plates. Reaction parameters were: 50 ˚C 4 min, 95 ˚C 5 min, followed by 40 cycles of 95 ˚C 15 sec, 60 ˚C 30 sec where fluorescence was measured each cycle. For each sample, ΔCt (Ct(*dps*)-

Ct(control)) was calculated. Then the average WT-dark ΔCt was subtracted from each ΔCt to generate a ΔΔCt, which is the same as $\log_2$(fold change) compared to WT-dark.

Protein expression and spectroscopy

E. litoralis lovK (gene locus Ga0102493_111685) and a variant of *lovK* in which the flavin adduct-forming cysteine was mutated to an alanine (C73A) were cloned into the pET28a expression plasmid; see **Table S6** for primers. Protein was expressed from these plasmids in *E. coli* BL21(DE3). Briefly, 2 liters of LB (containing 50 μg ml^{-1} kanamycin) was inoculated with a 100 ml of an overnight culture. The culture was grown at 37°C / 220 rpm to OD_{600} ≈0.8 and induced with 0.5 mM IPTG for four hours before cells were pelleted by centrifugation. For purification, the cell pellet was resuspended in 50 ml of resuspension buffer (10 mM Tris (pH 7.6), 150 mM NaCl, 10 mM imidizole) and lysed by two passages through a LV-1 microfluidizer. The lysate was then clarified by centrifugation (15 minutes at 35,000 x g) and applied to 2.5 ml of Ni-NTA resin on a gravity column. The resin had been pre-equilibrated with lysis buffer. The resin was washed with 5 column volumes (CV) of resuspension buffer, followed by a step gradient of buffer with increasing concentrations of imidazole. Specifically, the column was washed with 12 ml resuspension buffer containing 75 mM imidazole, 4 ml buffer with 200 mM imidazole, and finally 5 ml buffer with 500 mM imidazole for elution. The purity of the different fractions was assessed on a 12% SDS-PAGE gel. Peak fractions were desalted with a Zeba spin column (ThermoFisher) with a 7000 kDa MWCO. The visible absorption spectrum of purified LovK and LovK(C73A) was measured in a Tecan Spark 20M. The "lit" state of LovK was generated by illuminating purified protein with a panel of 96 LED bulbs (3 mm bulbs, 430 nm peak wavelength) held 10-20 cm from the sample for 30 seconds (until the visible yellow color was bleached).

5.6. SUMMARY

Extracytoplasmic function (ECF) sigma factors are a major class of environmentally-responsive transcriptional regulators. In *Alphaproteobacteria* the ECF sigma factor, σ^{EcfG}, activates general stress response (GSR) transcription and protects cells from multiple stressors. A phosphorylation-dependent protein partner switching mechanism, involving HWE/HisKA_2-family histidine kinases, underlies σ^{EcfG} activation. The identity of these sensor kinases and the signals that regulate them remain largely uncharacterized. We have developed the aerobic anoxygenic photoheterotrophic (AAP) bacterium, *Erythrobacter litoralis* DSM 8509, as a comparative genetic model to investigate GSR regulation. Using this system, we sought to define the contribution of visible light and a photosensory HWE kinase, LovK, to GSR transcription. We identified three HWE kinase genes that collectively regulate GSR: *gsrK* and *lovK* are activators, while *gsrP* is a repressor. GSR transcription is higher in the dark than in light, and the opposing activities of *gsrK* and *gsrP* are sufficient to achieve light-dependent differential transcription. In the absence of *gsrK* and *gsrP*, *lovK* alone is sufficient to regulate GSR transcription in response to light. This regulation requires a photochemically active LOV domain in LovK. Our studies establish a role for visible light and HWE kinases in light-dependent regulation of GSR transcription in *E. litoralis*, an AAP species.

6. APPENDIX B – FIGURES

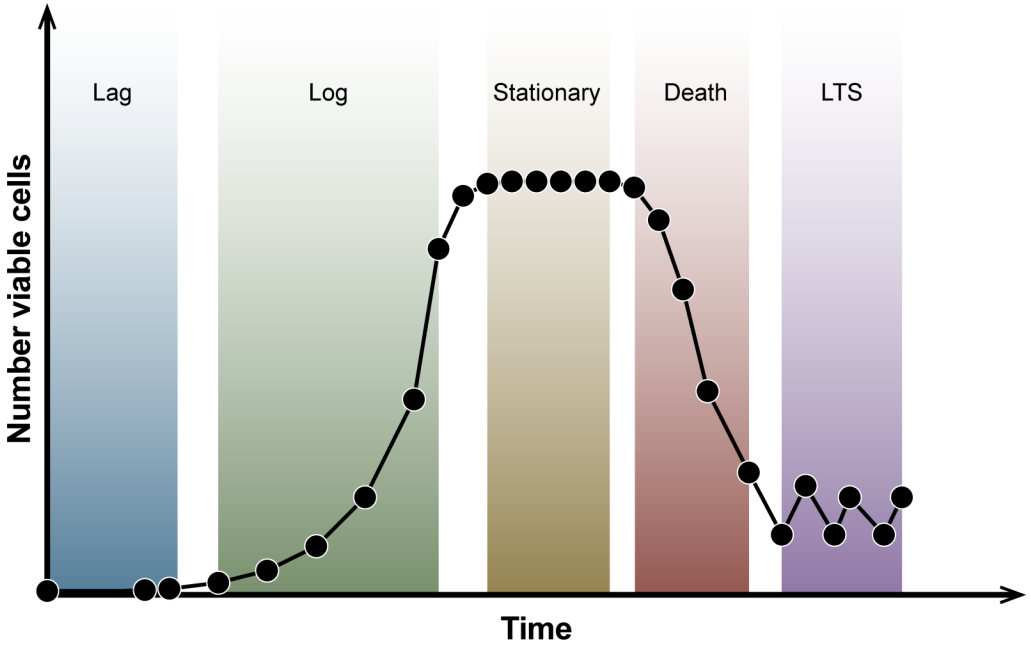


Figure 6.1 - Schematic representation of a growth curve
The initial lag phase (highlighted in **blue**) depends of the conditions the cells were in when the measurements commenced (frozen, stationary, different media, etc.). In log phase (or exponential phase, **green** shading) cells are in the best condition to replicate, and are doubling as fast as the conditions permit. Once nutrients run out (or other factors intervene, like density, stress, accumulation of toxic metabolites, etc.) cell growth slows down and they roll over, reaching a plateau, where the net number of cells does not increase (stationary phase, **ochre** shading). Stationary phase will persist for an amount of time that depends on a multitude of factors, including the specific organism in question, but will eventually lead to the death phase (**red** highlight) where cells are actively dying. At the end of the death phase, a small amount of persister cells will endure in the long-term stationary phase (LTS, **purple** highlight).

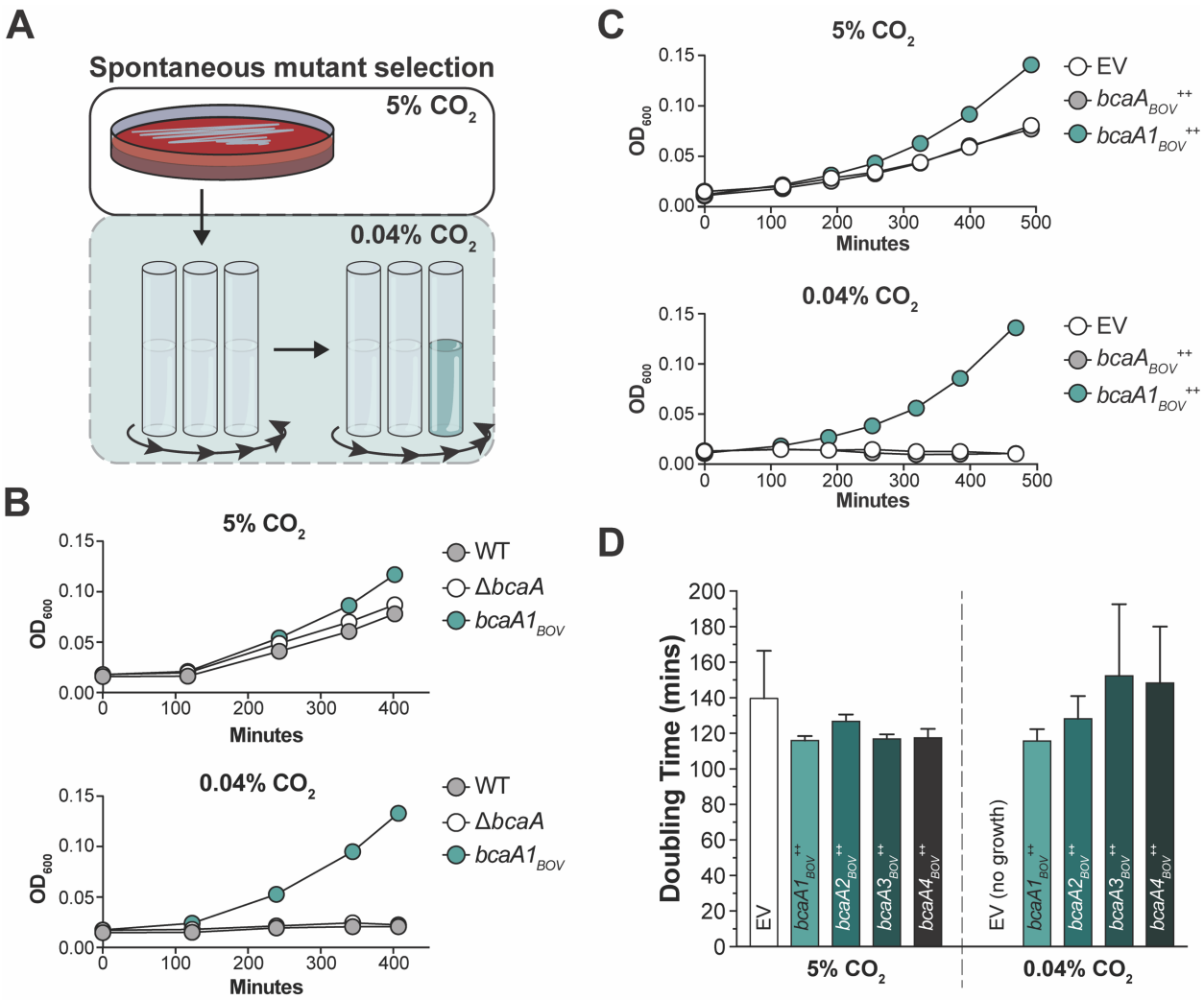


Figure 6.2 - Growth of *B. ovis* harboring *bcaA1-4*$_{BOV}$ alleles in an unsupplemented atmosphere

A) Cartoon illustrating the forward genetic selection for spontaneous mutants of *Brucella ovis* ATCC 25840 that grow without added CO_2. Cells were grown on plates at 37 °C with 5% CO_2 supplementation for 48 hrs then inoculated into BB and left in a shaking incubator at 37°C in standard atmospheric conditions (i.e. 0.04% CO_2). Growth was monitored by cell culture density measurements (OD_{600}) and individual colonies were isolated from tubes where growth was evident. **B)** Growth curve from three independent experiments of *Brucella ovis* ATCC 25840 (WT), Δ*bcaA* and *B. ovis bcaA1*$_{BOV}$ strains. Cells were grown either in 5% CO_2 (top) or 0.04% CO_2 (bottom). **C)** Growth curve from three independent experiments with *bcaA*$_{BOV}$ (*bcaA*$_{BOV}^{++}$) or *bcaA1*$_{BOV}$ (*bcaA1*$_{BOV}^{++}$) overexpressing strains. Cells were grown either in 5% CO_2 (top) or 0.04% CO_2 (bottom) after inducing expression from pSRK with 1mM IPTG. Strain carrying the empty vector (EV) was used as a control. **D)** Doubling time of *Brucella ovis* ATCC 25840 strains either carrying the pSRK empty vector plasmid (EV) or one of the four selected ("restored") alleles *bcaA1*$_{BOV}$, *bcaA2*$_{BOV}$, *bcaA3*$_{BOV}$ and *bcaA4*$_{BOV}$ (*bcaA1*$_{BOV}^{++}$, *bcaA2*$_{BOV}^{++}$, *bcaA3*$_{BOV}^{++}$ and *bcaA4*$_{BOV}^{++}$, respectively). Strains that did not grow are indicated (no growth). Experiment was

Continuation: Figure 6.2 – Growth of *B. ovis* harboring *bcaA1-4$_{BOV}$* alleles in an unsupplemented atmosphere

performed on separate days with technical replicates. Bar graphs indicate the standard deviation for 12 measurements per strain.

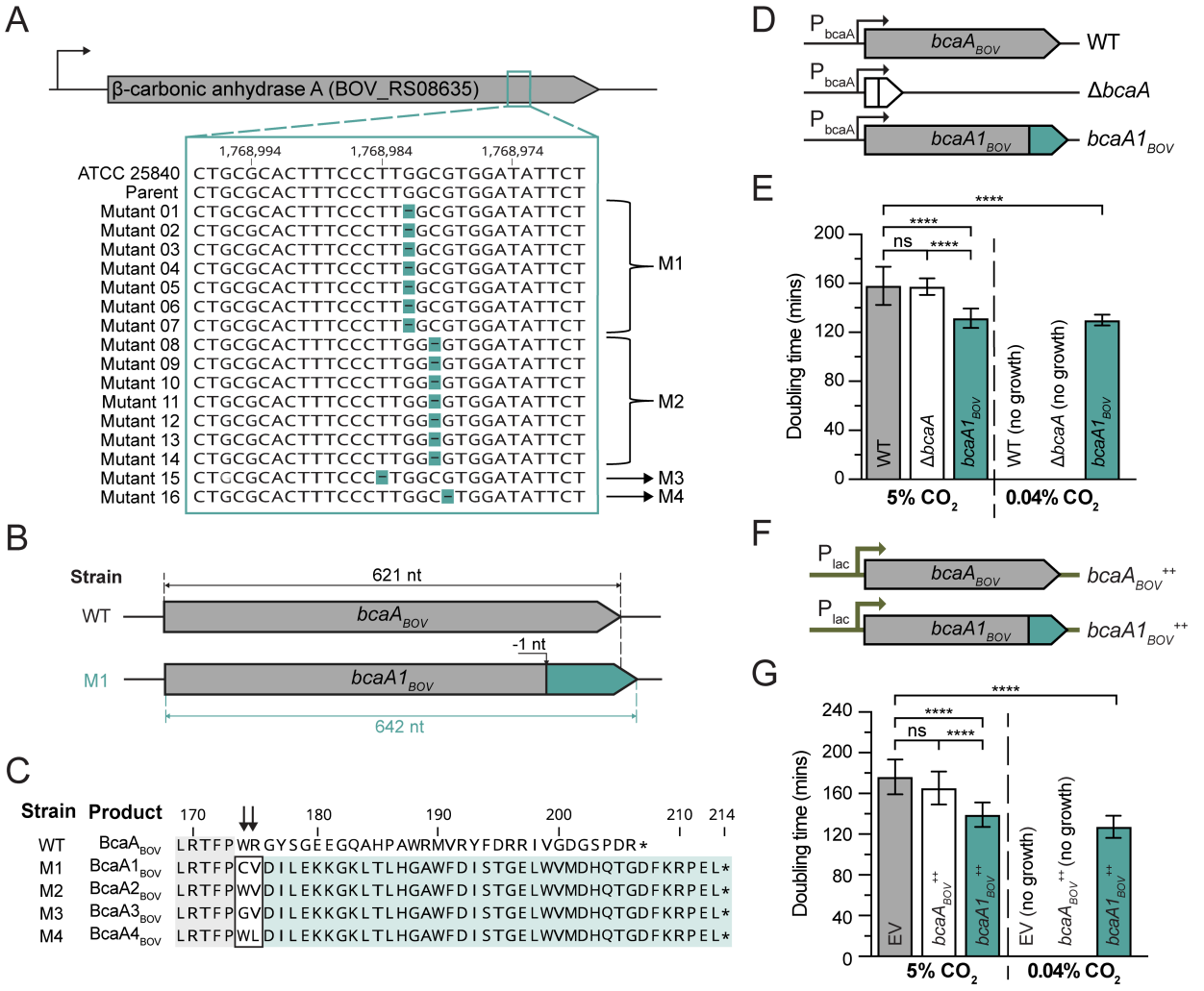
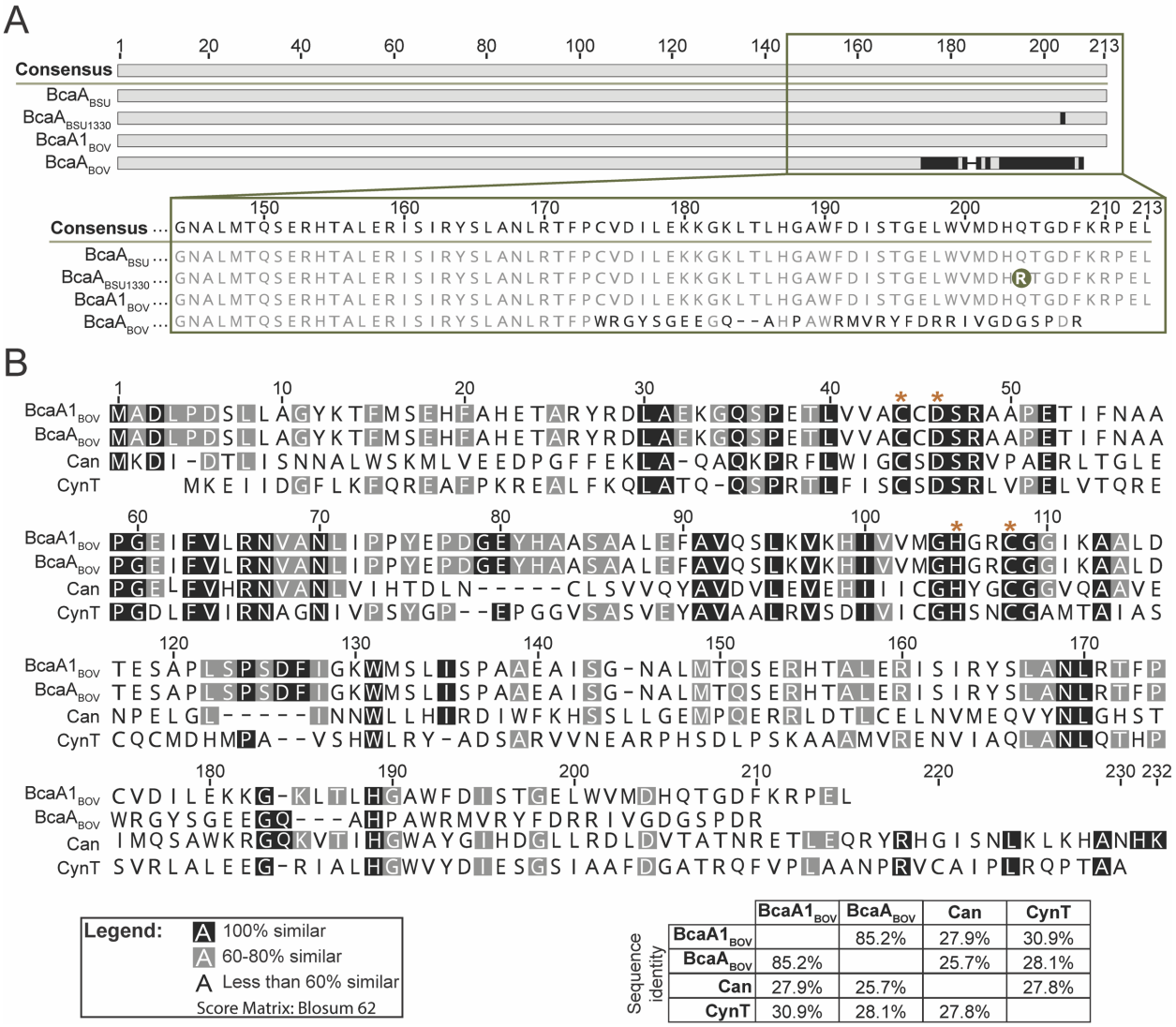


Figure 6.3 - A single nucleotide deletion at the 3' end of *bcaA$_{BOV}$* enables *B. ovis* growth without CO_2 supplementation

A) Alignment of a segment of the 3' end of *bcaA$_{BOV}$* (*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

Continuation: Figure 6.3 – A single nucleotide deletion at the end of the 3' end of the *bcaA*$_{BOV}$ enables *B. ovis* growth without CO_2 supplementation

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)** Bar graph of doubling times in minutes 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)** Bar graph of doubling times (in minutes) 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.



	BcaA1$_{BOV}$	BcaA$_{BOV}$	Can	CynT
BcaA1$_{BOV}$		85.2%	27.9%	30.9%
BcaA$_{BOV}$	85.2%		25.7%	28.1%
Can	27.9%	25.7%		27.8%
CynT	30.9%	28.1%	27.8%	

Figure 6.4 - Comparison of *B. suis* BcaA and *E. coli* β-carbonic anhydrases to BcaA$_{BOV}$ and BcaA1$_{BOV}$

A) (Top) Multiple alignment of *Brucella ovis* BcaA$_{BOV}$ against *Brucella suis* ATCC 23445 (BcaA$_{BSU}$) and *Brucella suis* 1330 (BcaA$_{BSU1330}$) homologs. The selected allele BcaA1$_{BOV}$, which enables growth of *B. ovis* without added CO_2, is included for comparison. Highlighted in black are the differences between each sequence and the genus-level consensus at top (see also Fig. 3 legend). (Bottom) Zoom-in on the C-terminal portion of the alignment that contains the *bcaA$_{BOV}$* frameshift present in wild-type *B. ovis*. A glutamine to arginine difference at position 204 of BcaA$_{BSU1330}$ is shaded in green. **B)** Multiple alignment of Escherichia coli MG1655 Can and CynT β-carbonic anhydrases with BcaA$_{BOV}$ and BcaA1$_{BOV}$. Orange asterisks show the conserved residues at the active site. Bottom right: table showing the amino acid identity between the proteins.

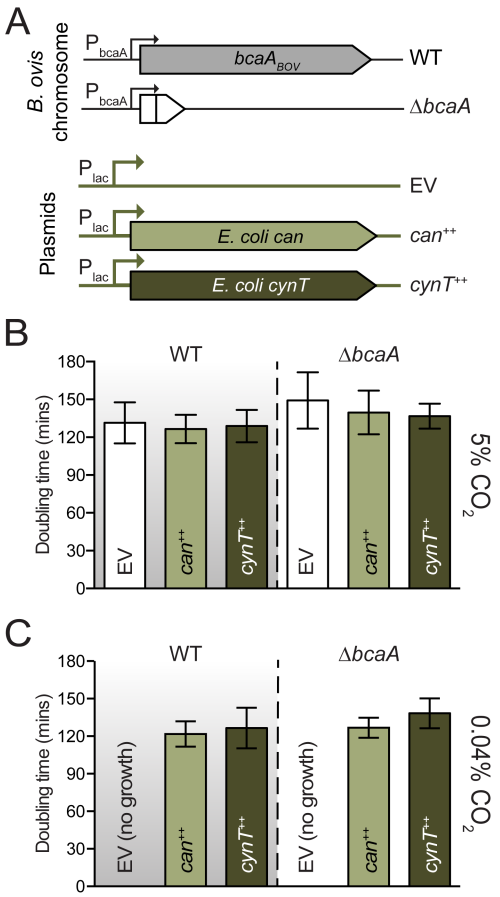


Figure 6.5 - Heterologous expression of two different *Escherichia coli* β-carbonic anhydrases enables growth of wild-type *B. ovis* ATCC 25840 without CO_2 supplementation

A) Schematic representation of RK2-derived plasmids (pSRK) transformed into either wild-type *B. ovis* ATCC 25840 (WT) or the in-frame *bcaA* deletion strain (Δ*bcaA*). 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. **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 at the top 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".

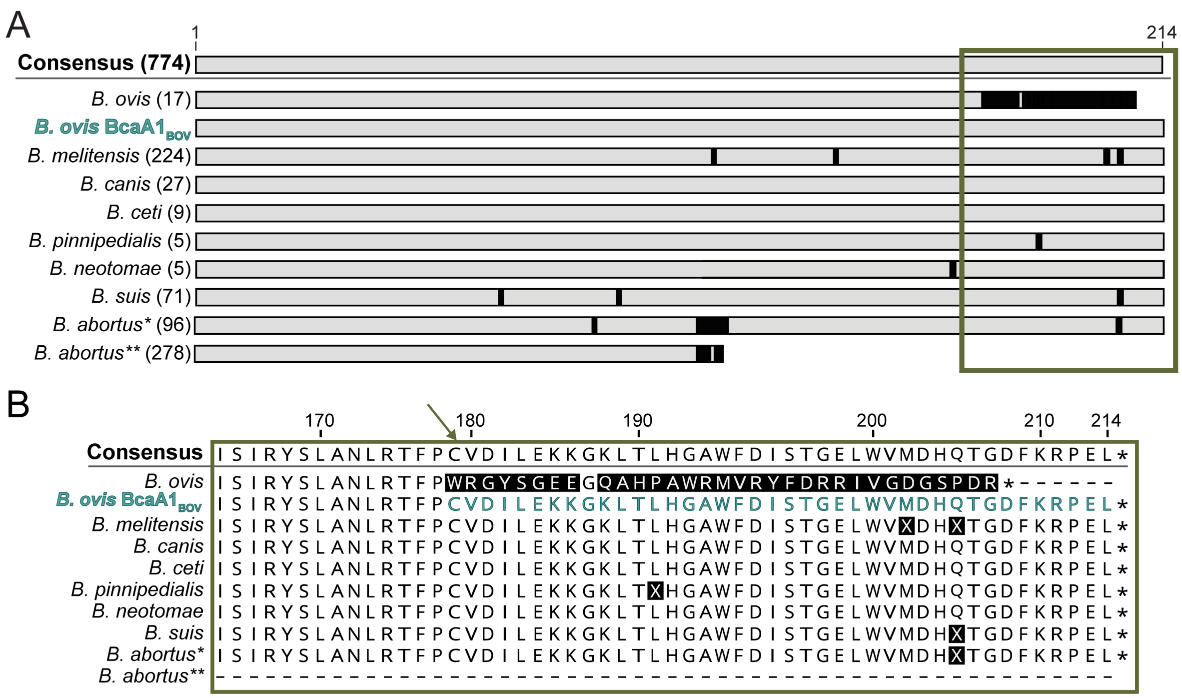


Figure 6.6 - Sequence polymorphisms in BcaA orthologs across the genus *Brucella*
A) Schematic of the BcaA protein sequence polymorphisms based on alignment of 774 BcaA orthologs from *Brucella* spp. available in the PATRIC database and found in **Data Set 1**, including *B. ovis* BcaA1$_{BOV}$ (**teal**; 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. 5** 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. 6** and **Results)**. 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**).

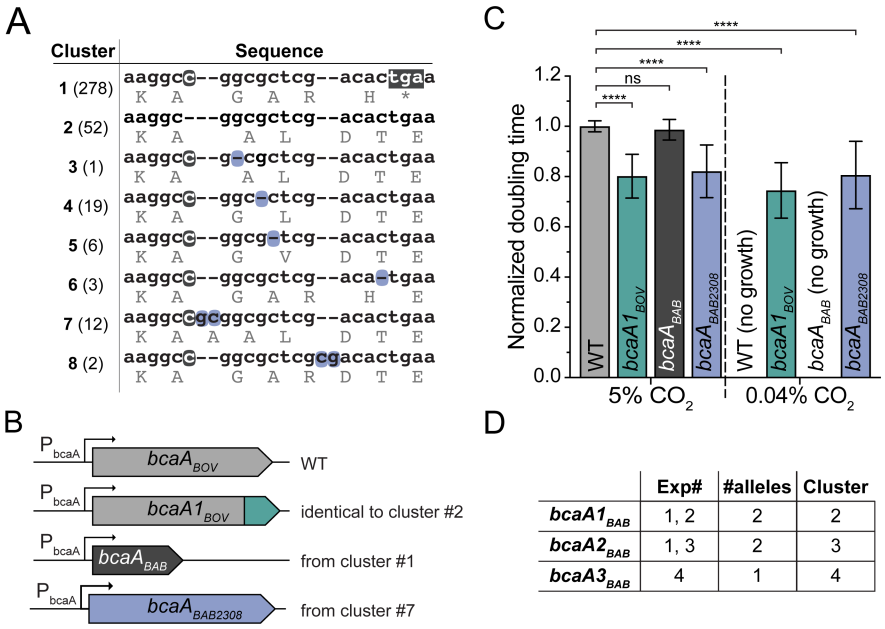


Figure 6.7 - Analysis of sequence and function of *B. abortus bcaA* alleles in *B. ovis* ATCC 25840

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.**

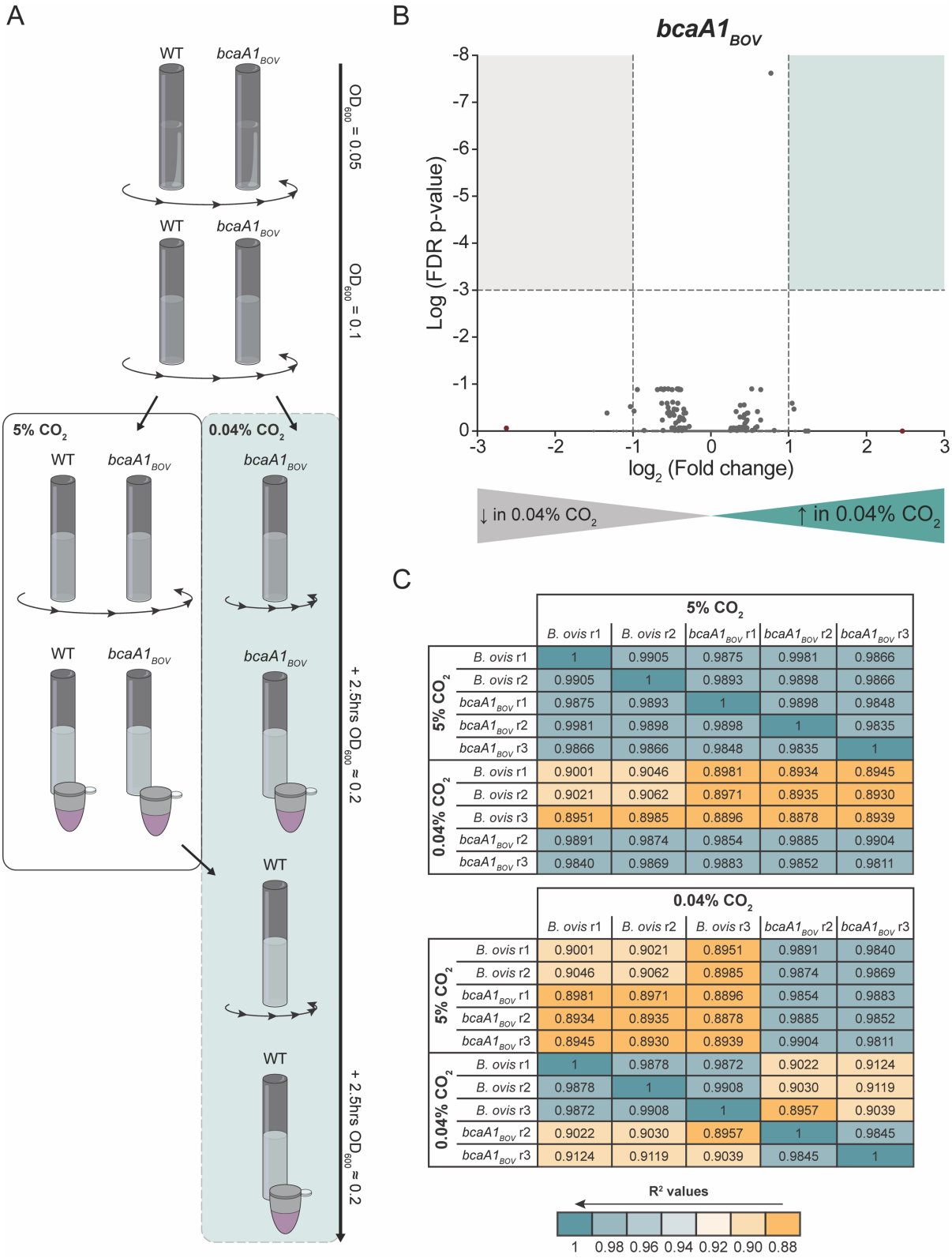


Figure 6.8 - RNA-seq experimental set up and measured gene expression changes in *B. ovis bcaA1*$_{BOV}$ upon CO_2 downshift

Continuation: Figure 6.8 - RNA-seq experimental set up and measured gene expression changes in *B. ovis bcaA1$_{BOV}$* upon CO_2 downshift

A) Schematic of sample treatment prior to RNA-seq. *B. ovis* wild-type and *bcaA1$_{BOV}$* strains were inoculated in BB at OD_{600} of 0.05 and grown with 5% CO_2 (6 tubes per strain, only one is shown for clarity). Once cells reached OD600 ~0.1, 3 tubes of *B. ovis bcaA1$_{BOV}$* cells were moved to roller in an air incubator (0.04% CO_2). Cells in all tubes were 5 then grown to OD_{600} ~0.2, and the three culture tubes of *Brucella ovis* ATCC 25840 (WT) in 5% CO_2 and all six B. ovis bcaA1BOV (*bcaA1$_{BOV}$*) cultures (3 from 5% CO_2 and 3 from 0.04% CO_2) were harvested at this point for RNA prep. Three remaining WT tubes were then moved to the air incubator (0.04% CO_2) where they were incubated for 2.5 hours before harvesting for RNA preparation. **B)** Volcano plot for genes differentially expressed in the *B. ovis bcaA1$_{BOV}$* strain upon downshift from 5% CO_2 to 0.04% CO_2. Analysis as described for wild-type *B. ovis* ATCC 25840 in Fig. 5B. No genes showed significant differential expression between these two CO_2 conditions (FDR < 0.001 and $\log_2$ fold change > |1|). See also Data Set 2. **C)** R2 values calculated when plotting $\log_{10}$ Counts Per Million (CPM) of *B. ovis* ATCC 25840 (WT) against *B. ovis bcaA1$_{BOV}$* in 0.04% or 5% CO_2. Replicates (r1-3) are indicated. Low R2 values are highlighted in a yellow gradient, high R2 values in a teal gradient, legend at the bottom.

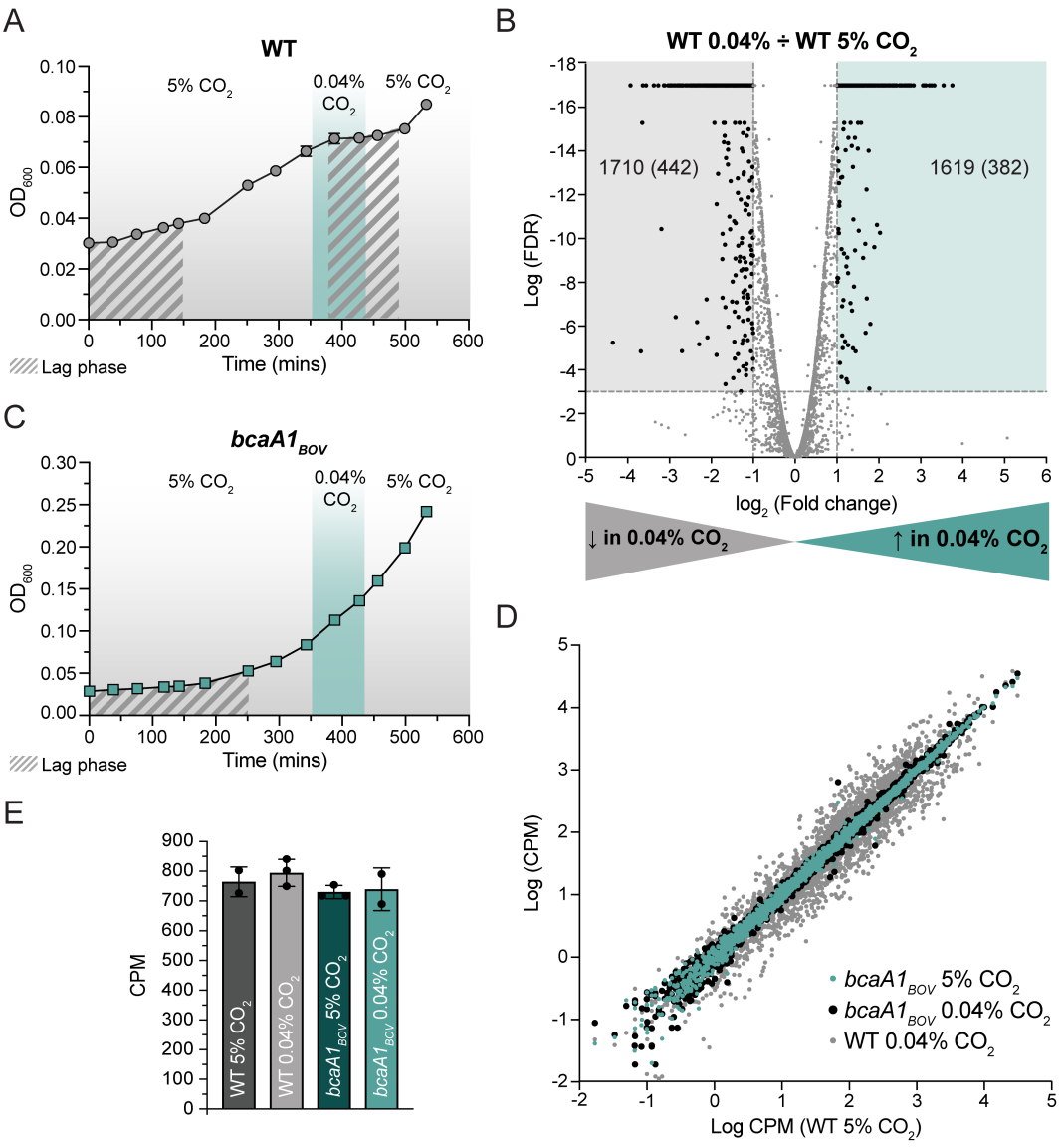


Figure 6.9 - Gene expression changes in wild-type *B. ovis* and *B. ovis bcaA1$_{BOV}$* upon CO_2 downshift from 5% to 0.04%

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

Continuation: Figure 6.9 - Gene expression changes in wild-type *B. ovis* and *B. ovis bcaA1$_{BOV}$* upon CO_2 downshift from 5% to 0.04%

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 CO_2 (**gray**). Each dot represents the mean CPM for a gene (see **Data Set 1**). R^2 values were calculated (see **Results** and **Fig. 7C**) implementing linear regression using GraphPad Prism 8. CPM 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.

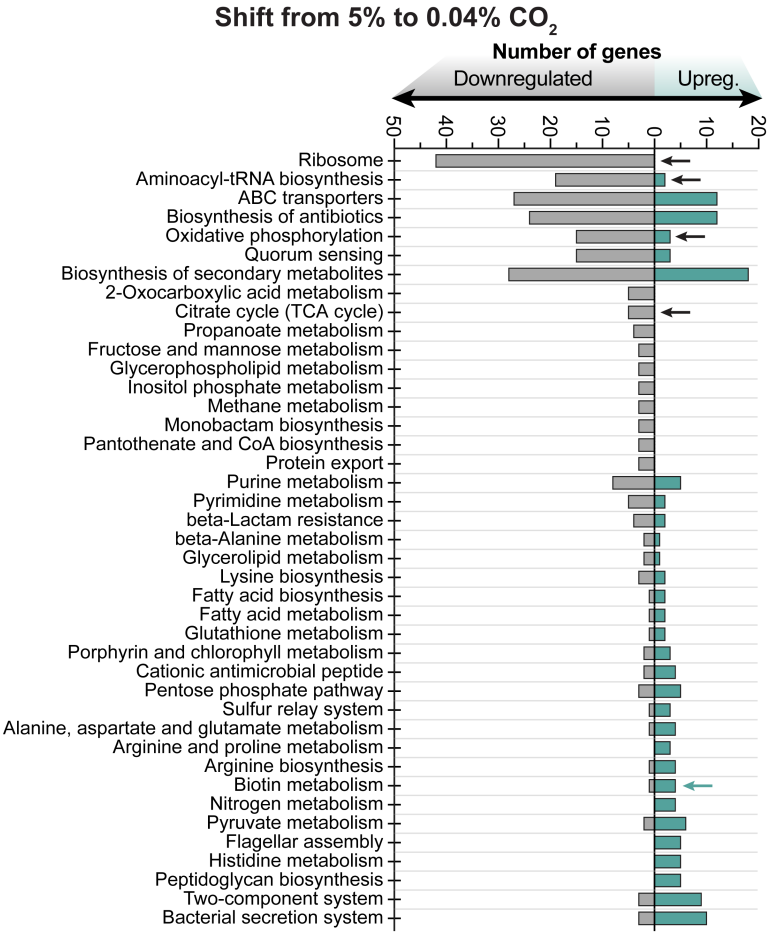


Figure 6.10 - KEGG pathway assignment of genes regulated in wild-type *B. ovis* in response to CO_2 downshift

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 those outlined in **Fig. 8B** legend. Pathways with less than 3 genes, or where the number of upregulated versus downregulated genes was not

Continuation: Figure 6.10 - KEGG pathway assignment of genes regulated in wild-type *B. ovis* in response to CO_2 downshift
significantly different (see **Materials and methods**), were removed from graph. Select KEGG categories associated with the stringent response are marked by **black arrows**. **Teal arrow** marks the "Biotin metabolism" pathway.

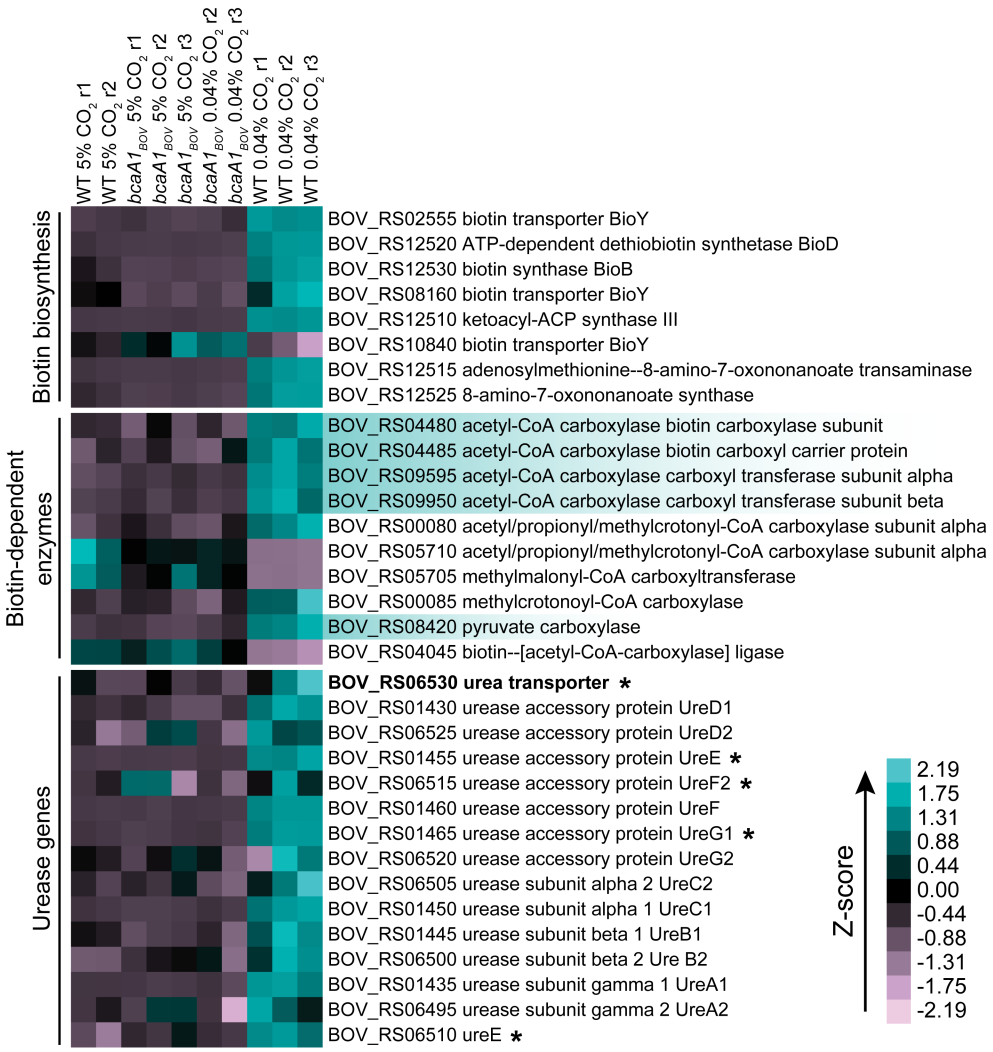


Figure 6.11 - Heat map representation of a subset of genes regulated by CO_2 downshift in wild-type *B. ovis*
B. ovis gene accession numbers and RefSeq annotations are shown to the right of the heatmap. Genes were grouped into three categories: biotin biosynthesis, biotin-dependent enzymes and urease 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 as 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**).

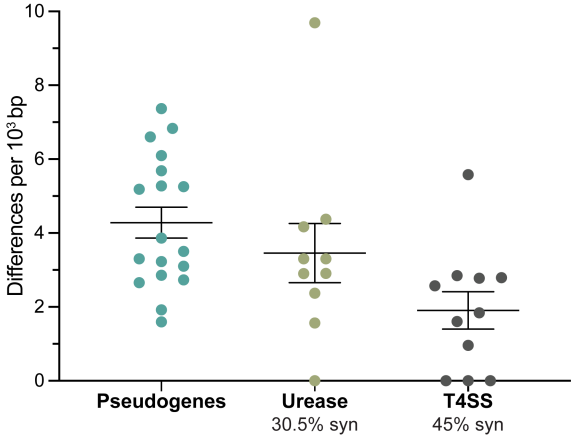


Figure 6.12 - Nucleotide differences between sets of functional genes and pseudogenes in *B. ovis* ATCC 25840 and their *B. abortus* ATCC 2308 orthologs normalized by gene length
Each dot represents a gene. Genes are cataloged in **Table 5**. Bars represent the mean and standard error of the mean for each group. Percentage of synonymous mutations in urease and T4SS genes are indicated.

Figure 6.13 - Fitness profile of *B. ovis* transposon insertion mutants as a function of growth phase

Heat map of fitness scores of 118 genes with t-like significance score ≥ |4|. Genes were grouped by hierarchical clustering. Each column is a point during the growth curve (OD_{600}, indicated at the

Continuation: Figure 6.13 - Fitness profile of *B. ovis* transposon insertion mutants as a function of growth phase
top of the heatmap). Each row is a gene. Gene locus tags and annotated functions are indicated on the left.

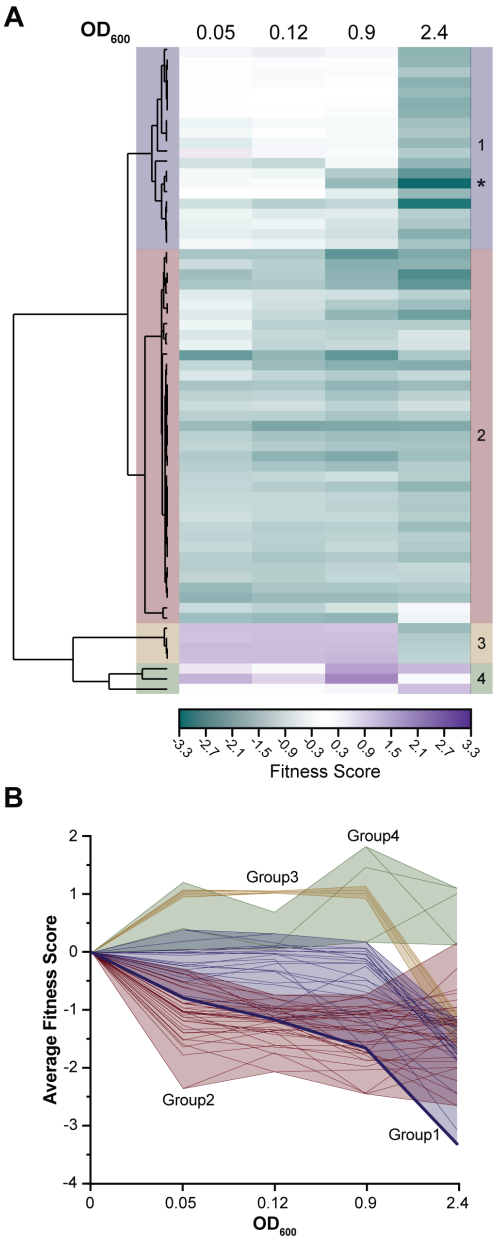
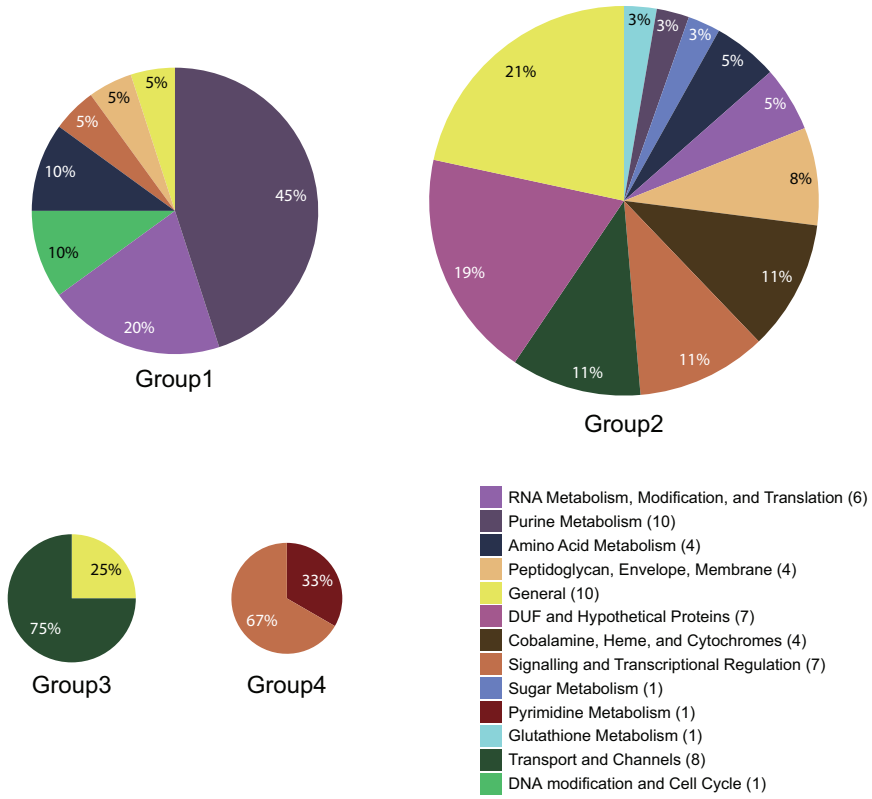


Figure 6.14 - Assessment of *B. ovis* mutant strain fitness as a function of growth phase identifies *cysE* as a determinant of stationary phase fitness
A) Heat map showing mean fitness scores (n = 3) of *B. ovis* mutant strains harboring barcoded transposon insertions in non-essential genes. Each row represents a gene, and each column is a point during the growth curve in BB (OD_{600}). Genes with a t-like significance score ≥|4| and fitness value ≥|1| in at least one timepoint are included in the heat map. Genes were hierarchically clustered (left), which yielded four main groups (right). The fitness profile of strains harboring RB

Continuation: Figure 6.14 - Assessment of *B. ovis* mutant strain fitness as a function of growth phase identifies *cysE* as a determinant of stationary phase fitness
Tn-*himar* insertions in *cysE* is marked with an asterisk on the right of the heat map. **B)** Alternative representation of data shown in **A**, where the average fitness score for each mutant (gene) is plotted as a function of timepoint. Individual lines represent genes, shaded area delimits the max and minimum values within that group. *cysE* is presented as a thick blue line. Group 1, blue; Group 2, red; Group 3, yellow; Group 4, green.

A



B

Annotated product	Locus ID	Average fitness score				Group
		0.05	0.12	0.9	2.4	
serine O-acetyltransferase	*BOV_06060*	- 0.18	- 0.11	- 1.65	- 3.31	1
AMP-nucleosidase	*BOV_03110*	- 0.79	- 1.16	- 0.99	- 3.06	1
bifunctional (p)ppGpp synthetase/guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase	*BOV_03230*	- 1.52	- 0.18	- 1.78	- 2.65	2
saccharopine dehydrogenase family protein	*BOV_01785*	- 0.2	- 0.32	- 1.22	- 2.45	1
MerR family DNA-binding transcriptional regulator	*BOV_05655*	- 1.33	- 1.29	- 1.72	- 2.43	2
alpha-beta hydrolase	*BOV_06975*	- 0.52	- 0.93	- 1.73	- 2.21	2

Figure 6.15 - Functional classification of *B. ovis* genes whose disruption significantly impacts fitness during growth in Brucella broth
A) Gene function pie chart of the four clusters of genes identified in **Fig. 2,** grouped based on predicted function. Area of each pie chart is proportional to the number of genes in that group.

Continuation: Figure 6.15 - Functional classification of *B. ovis* genes whose disruption significantly impacts fitness during growth in Brucella broth
DUF = Domain of Unknown Function. **B)** Subset of genes from **Fig. 2A** with fitness score values ≥ |2| in stationary phase.

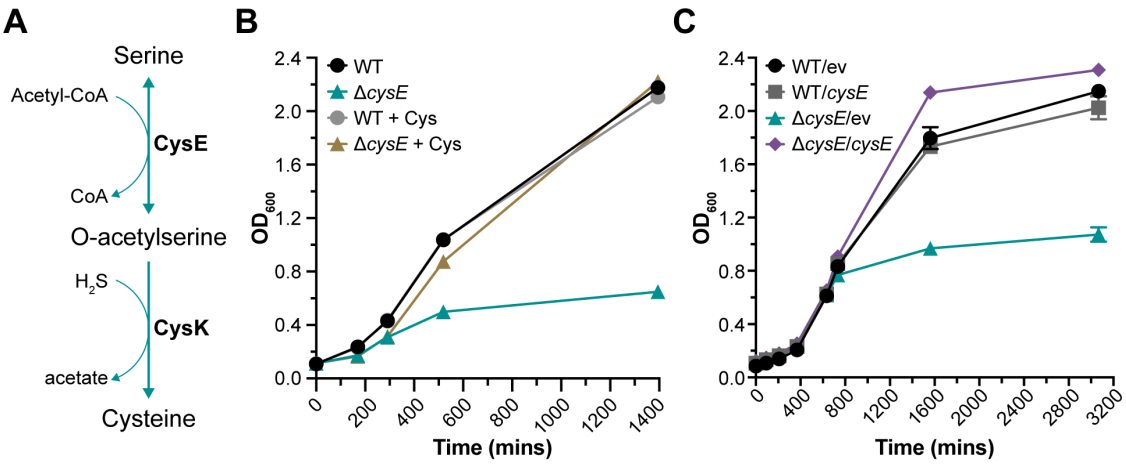


Figure 6.16 - Δ*cysE* enters stationary phase prematurely; this growth defect is rescued by addition of cysteine to the growth medium
A) Schematic of cysteine biosynthesis from serine. Teal arrows indicate cysteine biosynthesis enzymes annotated in the *Brucella ovis* genome (RefSeq accessions NC_009505 and NC_009504). Enzymes: CysE (*BOV_RS06060*, *cysE*, O-acetylserine transferase); CysK1 (*BOV_RS09280*, *cysK* , cysteine synthase A). **B)** Representative growth curves of wild type (WT, circles), and Δ*cysE* (triangles) with (gray and ochre, respectively) or without (black and teal, respectively) addition of 4 mM cysteine (Cys) to the growth medium. **C)** Representative growth curves of WT carrying the pSRK empty vector (EV, black circle) or pSRK-*cysE* (gray square), and Δ*cysE* carrying pSRK (teal triangle) or pSRK-*cysE* (purple diamond) in BB with 1 mM IPTG, 50 µg/ml Kan. Error bars represent standard deviation of technical replicates in representative experiments.

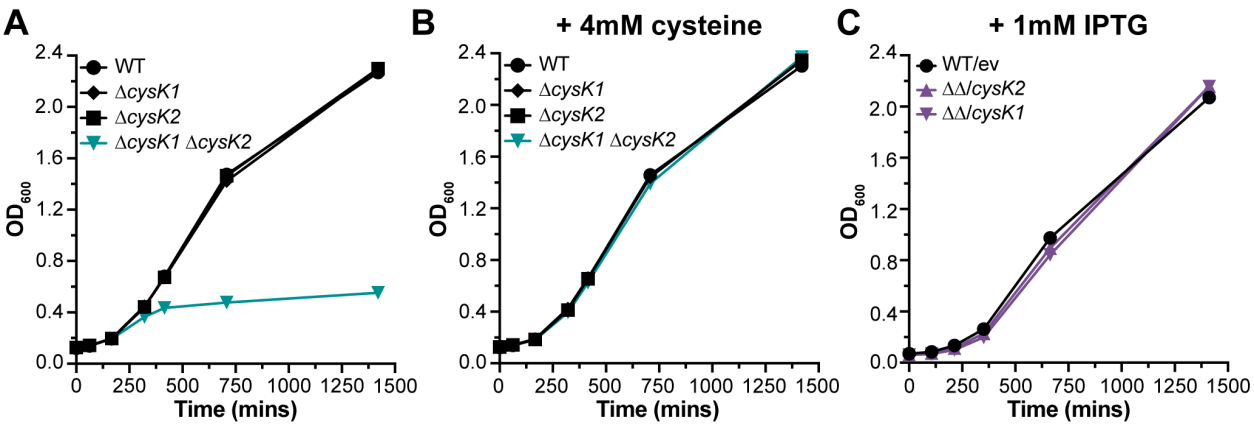


Figure 6.17 - The stationary phase phenotype of a Δ*cysK1* Δ*cysK2* double deletion phenocopies Δ*cysE* and is rescued by cysteine

A) Growth of wild type (black circles), Δ*cysK1* (black diamonds), Δ*cysK2* (black squares) and Δ*cysK1* Δ*cysK2* (teal triangles) in BB. **B)** Growth of the same four strains in **A** with 4 mM cysteine added to the broth. **C)** Growth curves of WT carrying the pSRK empty vector (black circles), and Δ*cysK1* Δ*cysK2* carrying either pSRK-*cysK1* or pSRK-*cysK2* (purple triangles) grown with 50 μg/ml Kan and 1 mM IPTG. Error bars represent standard deviation of technical replicates and may be smaller than symbol size. Growth curves were conducted at least three independent times. A representative curve is shown for each.

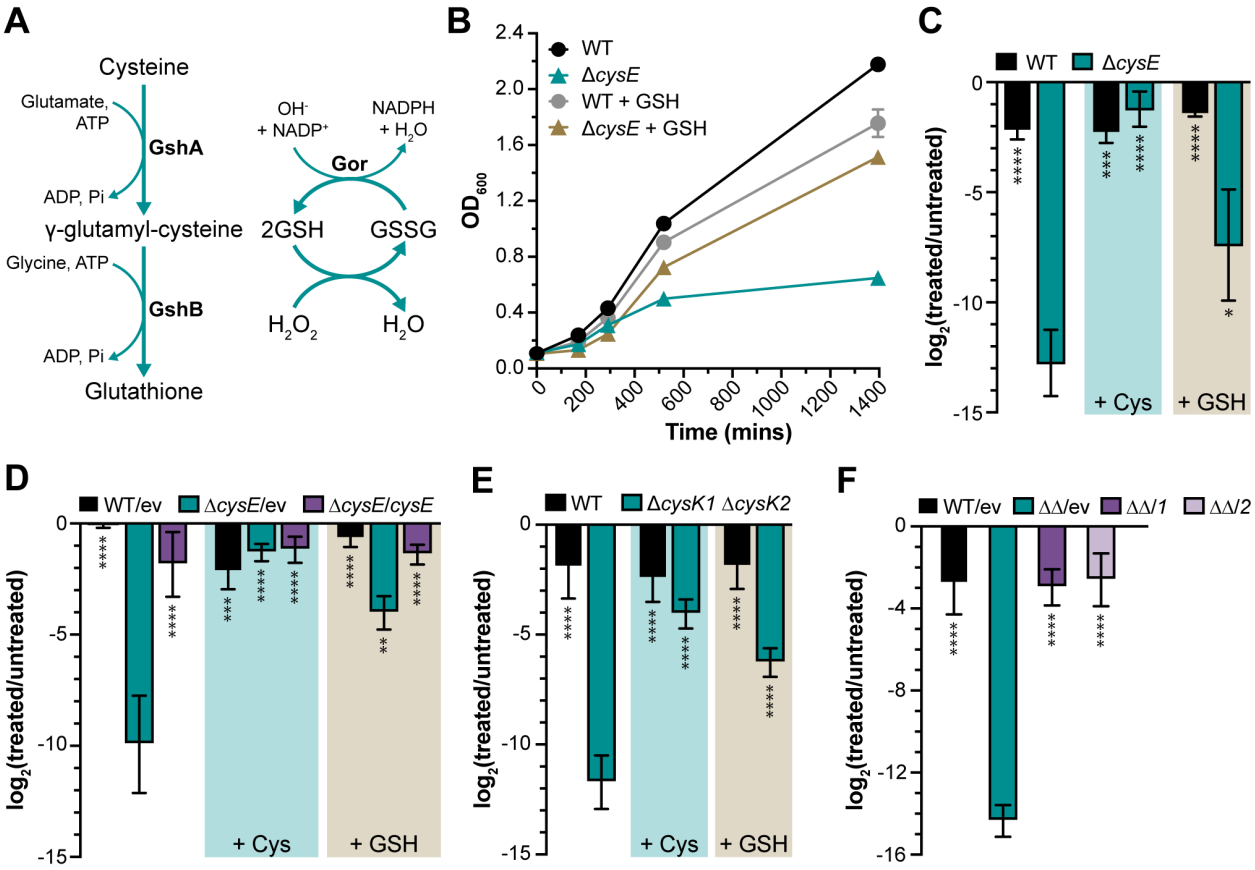


Figure 6.18 - *B. ovis* Δ*cysE* is sensitive to H_2O_2 treatment; Δ*cysE* growth defect and peroxide sensitivity is mitigated by glutathione

A) Schematic of glutathione metabolism. Enzymes annotated in the *Brucella ovis* genome (RefSeq accessions NC_009505 and NC_009504) are indicated in bold: GshA (*BOV_RS13935*, glutamate—cysteine ligase), GshB (*BOV_RS10075*, glutathione synthase), and Gor (*BOV_RS04850*, glutathione disulfide–reductase). GSH (glutathione, reduced state); GSSH (glutathione disulfide, oxidized state). **B)** Growth of wild type (circles) or Δ*cysE* (triangles) in BB with (gray and ochre, respectively) or without (black and teal, respectively) 4 mM GSH added to the medium. Error bars represent standard deviations and may be smaller than symbol size. **C)** Hydrogen peroxide survival assay showing the $\log_2$ ratio of CFU of treated (20 mM H_2O_2) versus untreated (mock PBS) cultures. Black bars represent wild type and teal bars represent Δ*cysE* strains. Addition of either 4 mM cysteine (+ Cys) or 4 mM GSH (+ GSH) to the medium is indicated by the shaded boxes. GSH and cysteine was washed away from the culture prior to peroxide treatment. **D)** Same assay as in **C** but with strains carrying plasmids to test genetic complementation. Strains were treated with 15 mM H_2O_2. Black bars represent WT carrying an empty vector (WT/ pSRK-EV), teal bars represent Δ*cysE* carrying an empty vector (Δ*cysE* / pSRK-EV), and purple bars represent complemented Δ*cysE* (Δ*cysE* / pSRK-*cysE*). p-values: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$, **** $p < 0.0001$, calculated using one-way ANOVA (followed by Dunnett's multiple comparison test, to Δ*cysE* in **C** or Δ*cysE*/ev in **D**). Hydrogen peroxide survival assay as in **C**, comparing wild type *B. ovis* (black bars) to the Δ*cysK1* Δ*cysK2* double deletion strain (teal bars) **F)** Same as in **D** but with strains carrying plasmids to test genetic complementation. Black bars represent *B. ovis* carrying the empty vector pSRK-EV (WT/EV), teal

Continuation: Figure 6.18 - *B. ovis* Δ*cysE* is sensitive to H_2O_2 treatment; Δ*cysE* growth defect and peroxide sensitivity is mitigated by glutathione
bars represent Δ*cysK1* Δ*cysK1* with pSRK-EV (ΔΔ/*ev*); and purple bars represent the Δ*cysK1* Δ*cysK2* strain carrying either pSRK-*cysK1* (ΔΔ/*1*, dark purple) or pSRK-*cysK2* (ΔΔ/*2*, light purple). Error bars represent standard error of the mean for 3 or 4 independent experiments.

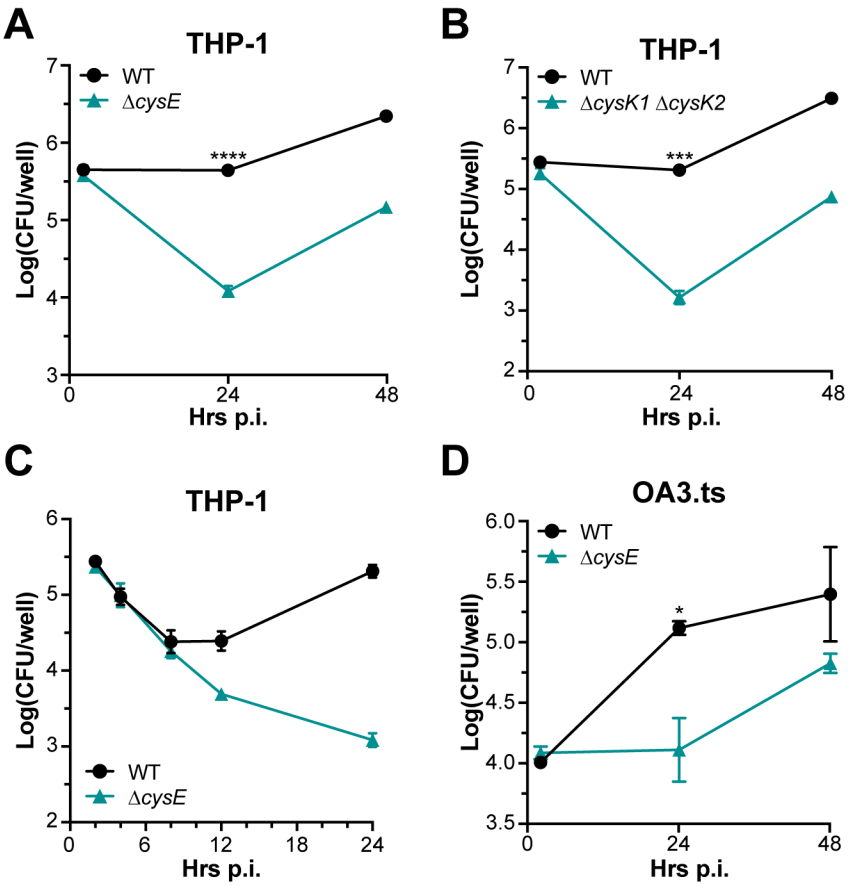


Figure 6.19 - *B. ovis* Δ*cysE* has reduced fitness in the intracellular niche of human macrophage-like cells and an ovine testis epithelial cell line
Log_{10} colony forming units (CFU) per well of WT (black circles) or Δ*cysE* (teal triangles) isolated from infected THP-1 (**A** and **C**) or OA3.ts mammalian cells (**D**) enumerated at 2, 24 and 48 hours (hrs) post infection (p.i.) (**A** and **D**) or at 2, 4, 8, 12, and 24 hrs p.i. (**C**). **B)** Log_{10} CFU/well per well of WT (black circles) or Δ*cysK1* Δ*cysK2* (teal triangles) strains isolated from infected THP-1 enumerated at 2, 24 and 48 hrs p.i. p-value comparing recovered *B. ovis* 24 hrs p.i. were calculated using an unpaired t-test **A, B** and **D**. * $p < 0.05$; *** $p < 0.001$, **** $p < 0.0001$. Infections were repeated 3 -5 independent times. Error bars represent standard error within the representative experiment.

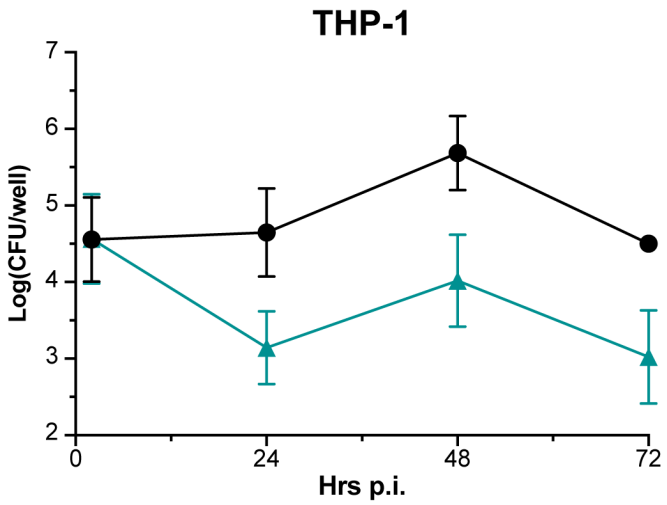


Figure 6.20 - Recovered CFUs decrease at 72 hrs post infection
Colony forming units of wild type *B. ovis* (WT, black circles) and Δ*cysE* (teal triangles) cells recovered from infected THP-1 cells at 2, 24, 48, and 72 hours post infection (Hrs p.i.). Error bars represent SEM for three independent experiments.

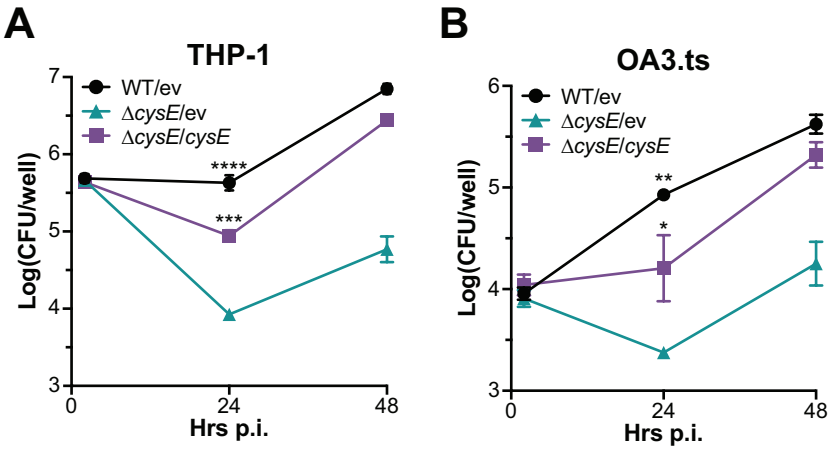


Figure 6.21 - Genetic complementation of THP-1 and OA3.ts infection
Log_{10} colony forming units (CFU) per well of WT/pSRK-EV (WT/ev, black circles), Δ*cysE*/pSRK-EV (Δ*cysE*/ev teal triangles), and Δ*cysE*/pSRK-cysE (Δ*cysE*/*cysE*, purple diamonds) isolated from infected THP-1 (**A**) or OA3.ts mammalian cells (**B**). Significance at the 24 hr time point was assessed using one-way ANOVA (Tukey's multiple comparison test). Asterisks indicate p-values compared to Δ*cysE* or Δ*cysE*/ev: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$, **** $p < 0.0001$. Infections were repeated 3-5 independent times. Error bars represent standard error within the representative experiment.

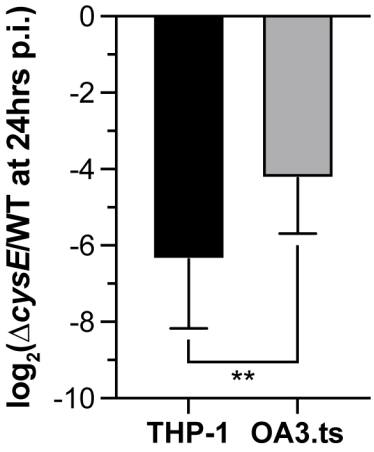


Figure 6.22 - *B. ovis* Δ*cysE* is more attenuated in THP-1 human macrophage-like cells than in ovine testis OA3.ts cells
Ratio of recovered CFUs of *B. ovis* Δ*cysE* relative to WT at 24 hrs post infection (p.i.) in macrophage-like human THP-1 cells and in the sheep testis epithelial cell line, OA3.ts. p-value was calculated using a two-tailed unpaired t-test, ** $p<0.01$. Error bars represent SD of four independent experiments in each cell line.

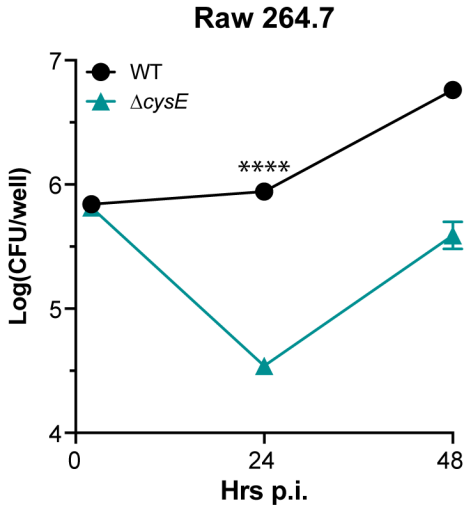


Figure 6.23 - *B. ovis* Δ*cysE* strain is attenuated in Raw 264.7 macrophages
Colony forming units of wild type (black circles) and Δ*cysE* (teal triangles) *B. ovis* cells recovered after infection of Raw 264.7 murine macrophages. p-value calculated using an unpaired t-test for the 24 hr time point. Error bars represent SEM for four independent repetitions of the experiment.

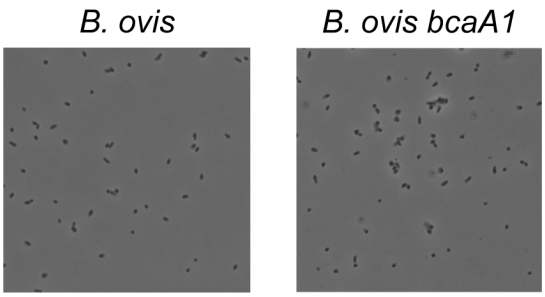
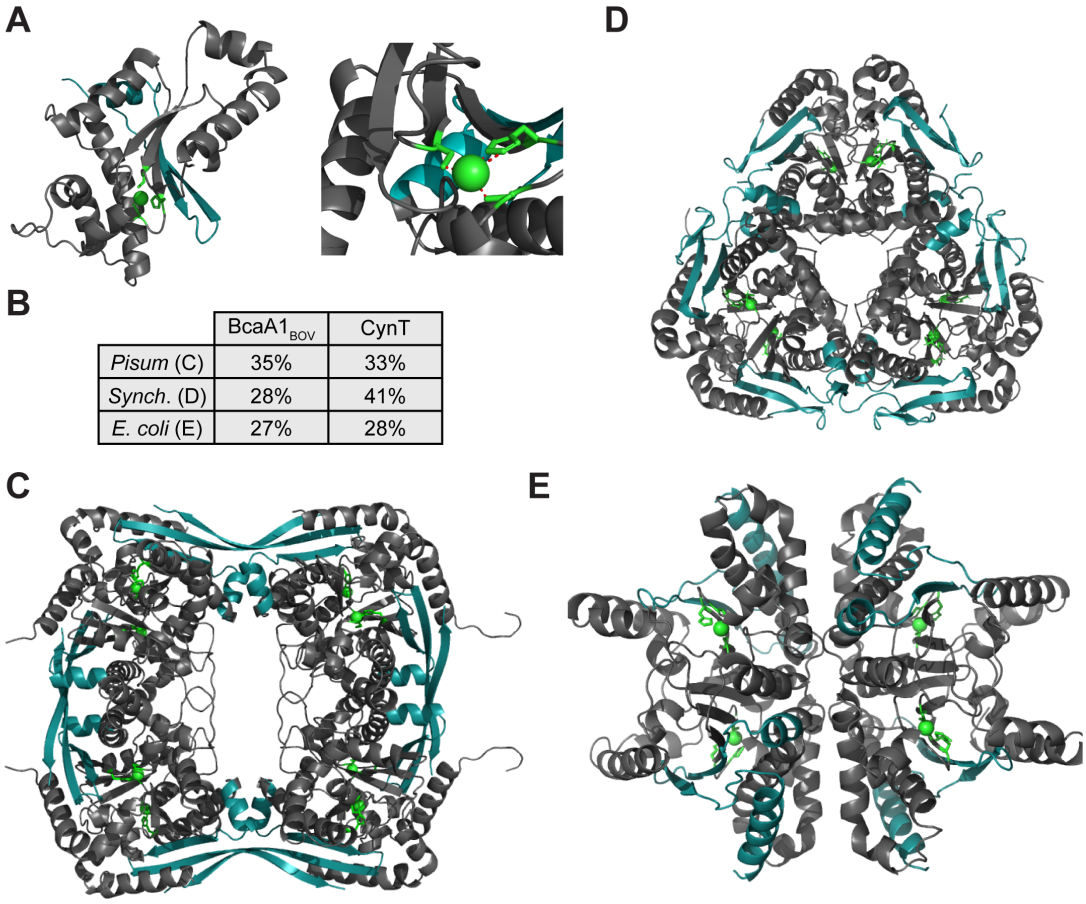


Figure 6.24 - *Brucella ovis* and *B. ovis bcaA1* have no difference in morphology
Contrast microscopy images showing wild type *B. ovis* (**left**) and *B. ovis bcaA1* (**right**) picked from a 48 h-old plate grown in 5% CO_2 supplementation. See **Materials and methods** from (244) (*B. ovis* eipA *depletion assays*) for additional information.



	$BcaA1_{BOV}$	CynT
Pisum (C)	35%	33%
Synch. (D)	28%	41%
E. coli (E)	27%	28%

Figure 6.25 - Modelling of BcaA structure in *B. ovis* and comparisons to experimental structures of other carbonic anhydrases
A) *Left*: modelling of $BcaA1_{BOV}$ based on the structure of β-carbonic anhydrase from *Pisum sativum*. In teal are the regions affected by the inactivating frameshift, in green the three amino acids that coordinate to the $Zinc^{2+}$ ion at the active site, detailed on the *right*. Models of BcaA1 monomer were built as follows. Amino acid sequences were implemented in XtalPred (http://ffas.burnham.org/XtalPred-cgi/xtal.pl). Search of PDB homologs yielded 53 sequences.

Continuation: Figure 6.25 - Modelling of BcaA structure in *B. ovis* and comparisons to experimental structures of other carbonic anhydrases
The most similar sequence (35% identical) for BcaA1 was from *Pisum Sativum* (PDB: 1ekj). The PDB file of the pea β-carbonic anhydrase sequence was fitted to adapt the BcaA1 sequence using Phyre[2] {Kelley, 2015 #210} and Swiss-Model {Waterhouse, 2018 #211}. The resulting PDB file was imported in PyMOL (the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) for graphic rendering. **B)** Table showing the similarities between the carbonic anhydrases in *Pisum sativum* (*Pisum*), *Synechocystis* sp. (strain PCC 6803 / Kazusa) (*Synch*.), and *Escherichia coli* (*E. coli*) to *B. ovis* BcaA1 and *E. coli* CynT. The structures for the *P. sativum*, *Synechocystis* (PDB: 5swc), and *E. coli* Can (PDB:1t75) carbonic anhydrases are shown in **C**, **D**, and **E**, respectively. The metal ions and coordinated amino acids are shown in green, the ribbons corresponding to the regions that would be affected by the frameshift inactivating $BcaA1_{BOV}$ are shown in teal.

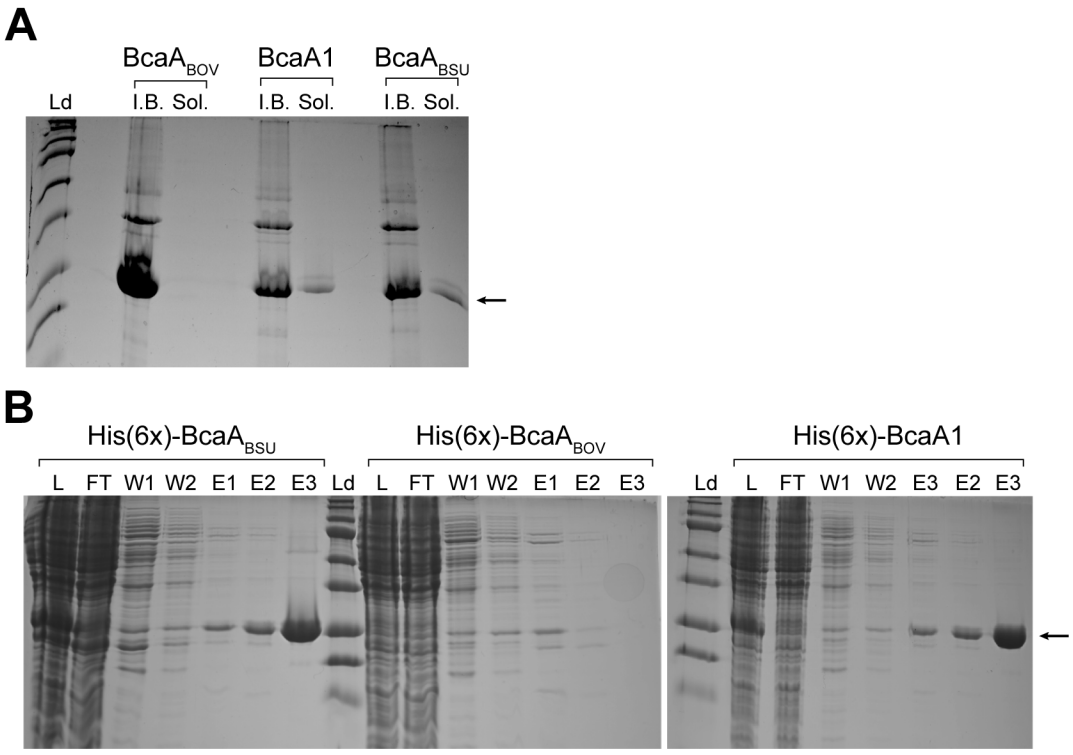


Figure 6.26 - $BcaA_{BOV}$ does not purify in a soluble form and is mainly found in inclusion bodies
E. coli Rosetta strains carrying pET28a-Hix(6x)-*bcaA$_{BOV}$*, pET28a-His(6x)-*bcaA1*, and pET28a-His(6x)-*bcaA$_{BSU}$* were inoculated in LB and grown at 37 °C, until $OD_{600} \approx 0.6$ was reached. Expression of N-terminally His-tagged proteins was induced with 1 mM IPTG for 4-6 hours. Cells were centrifuged and pellet was collected, resuspended in 50 ml of Binding buffer (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 10 mM Imidazole) supplemented with DNAse A (1 µl/ml), PMSF (4 µl/ml) and half a tablet of complete protease inhibitor mix (Roche Applied Science), and lysed using a microfluidizer (Microfluidics LV1). The lysate was clarified by centrifugation. The soluble fraction (Sol.) and insoluble fraction (inclusion bodies, I.B.) were run on a 14% SDS-page gel (A). Purification of the His-tagged proteins was achieved through nickel affinity chromatography (nitrilotriacetic acid resin). The lysate (L) was run through the column, and the flow through (FT) was collected as well. The two wash steps (W1 and W2) were done using Wash Buffer (10 mM

Continuation: Figure 6.26 – BcaA$_{BOV}$ does not purify in a soluble form and is mainly found in inclusion bodies
Tris-HCl pH 7.4, 150 mM NaCl, 75 mM imidazole) and the bound proteins eluted twice (E1 and E2) in elution buffer 1 (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 200 mM imidazole) and a third time (E3) in elution buffer 2 (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 500 mM imidazole). Samples from the various purification steps were collected and run on a 14% SDS-page gel (**B**).

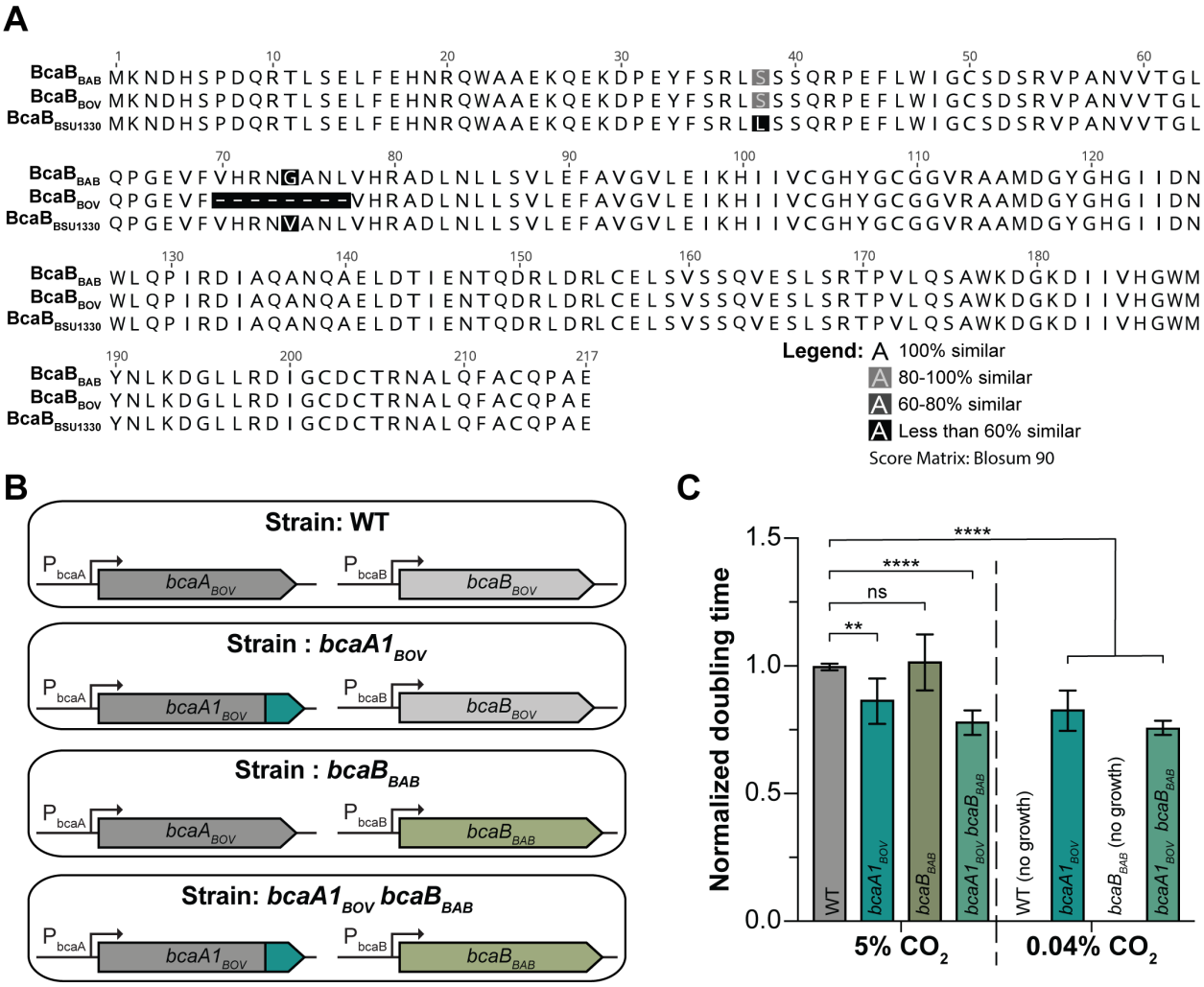


Figure 6.27 - BcaB from *B. abortus* cannot rescue *B. ovis* growth in 0.04% CO_2
A) Amino acid alignment comparison between BcaB from *Brucella abortus* ATCC 2308 (BcaB$_{BAB}$), *Brucella ovis* ATCC 25840 (BcaB$_{BOV}$), and *Brucella suis* 1330 (BcaB$_{BSU1330}$). Amino acids that do not match the consensus (based on Blosum 90 score matrix) are indicated in varying shades of gray as explained by the legend in the bottom right corner. B) Strains built for experiment in C: WT *B. ovis* (WT), *B. ovis bcaA1$_{BOV}$* (*bcaA1$_{BOV}$*), *B. ovis* or B. ovis bcaA1 carrying the *B. abortus bcaB* allele replacing *bcaB$_{BOV}$* at the native locus (*bcaB$_{BAB}$* and *bcaA1$_{BOV}$ bcaB$_{BAB}$*, respectively). C) Doubling time normalized to WT *B. ovis* of strains grown in atmospheric Continuation: Figure 6.27 – conditions (0.04% CO_2, *right*) or supplied with 5% CO_2 (*left*). p-value was calculated using one way Anova and Tukey's multiple comparison test, ns: non-significant, *

Continuation: Figure 6.27 – BcaB from B. abortus cannot rescue B. ovis growth in 0.04% CO_2
p-value < 0.05, ** p- value < 0.01, **** p-value < 0.0001. See **2.5 Materials and methods** for additional information on bacterial strains and growth conditions, plasmid and strain construction, growth curves, and sequence alignments.

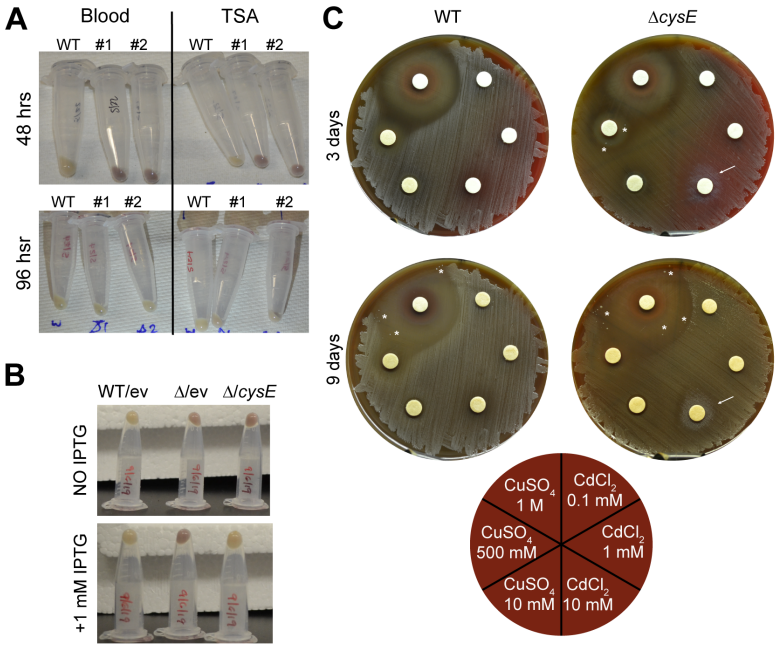


Figure 6.28 - Deletion of *cysE* leads to differences in coloring and tolerance to different metals
A) Wild type *Brucella ovis* (WT), *B. ovis* Δ*cysE* clone #1 (#1), or *B. ovis* Δ*cysE* clone #2 (#2) cells were scraped off of a 48 hour-old (top) or 96 hour-old (bottom) plain TSA (TSA) or TSA + 5% sheep blood plate (Blood). **B)** *B. ovis* carrying pSRK-EV (EV) or *B. ovis* Δ*cysE* carrying pSRK-EV (Δ/ev) or pSRK-cysE (Δ/*cysE*) cells scraped off of TSA + 5% sheep blood plates supplied with 50 µg/ml of kanamycin and 1 mM IPTG (**bottom**) or no IPTG (**top**). **C)** Disc diffusion assay plates were sterile cotton discs were imbued with different concentrations of either CdCl2 or $CuSO_4$ (bottom cartoon) and placed on lawns of *B. ovis* (WT) or *B. ovis* Δ*cysE* (Δ*cysE*). Pictures were taken after 3 days of spreading and 9 days. Asterisks indicate single colonies of $CuSO_4$ resistant cells and white arrows highlight metallic shade around the $CdCl_2$ 10 mM disc on lawns of *B. ovis* Δ*cysE* cells. See **3.5 Materials and methods** for details on strain building and additional information.

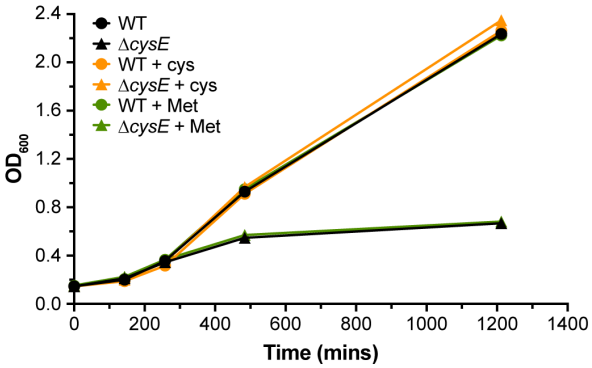


Figure 6.29 - Supplementation of methionine does not restore Δ*cysE* growth to wild type levels

Growth of WT *B. ovis* (WT, **circles**) or *B. ovis* Δ*cysE* mutant strain (Δ*cysE*, **triangles**) with cysteine (**orange**), methionine (**green**), or without supplemental additions to the media (**black**). See **3.5 Materials and methods** for details on strain building, growth curves, and additional information.

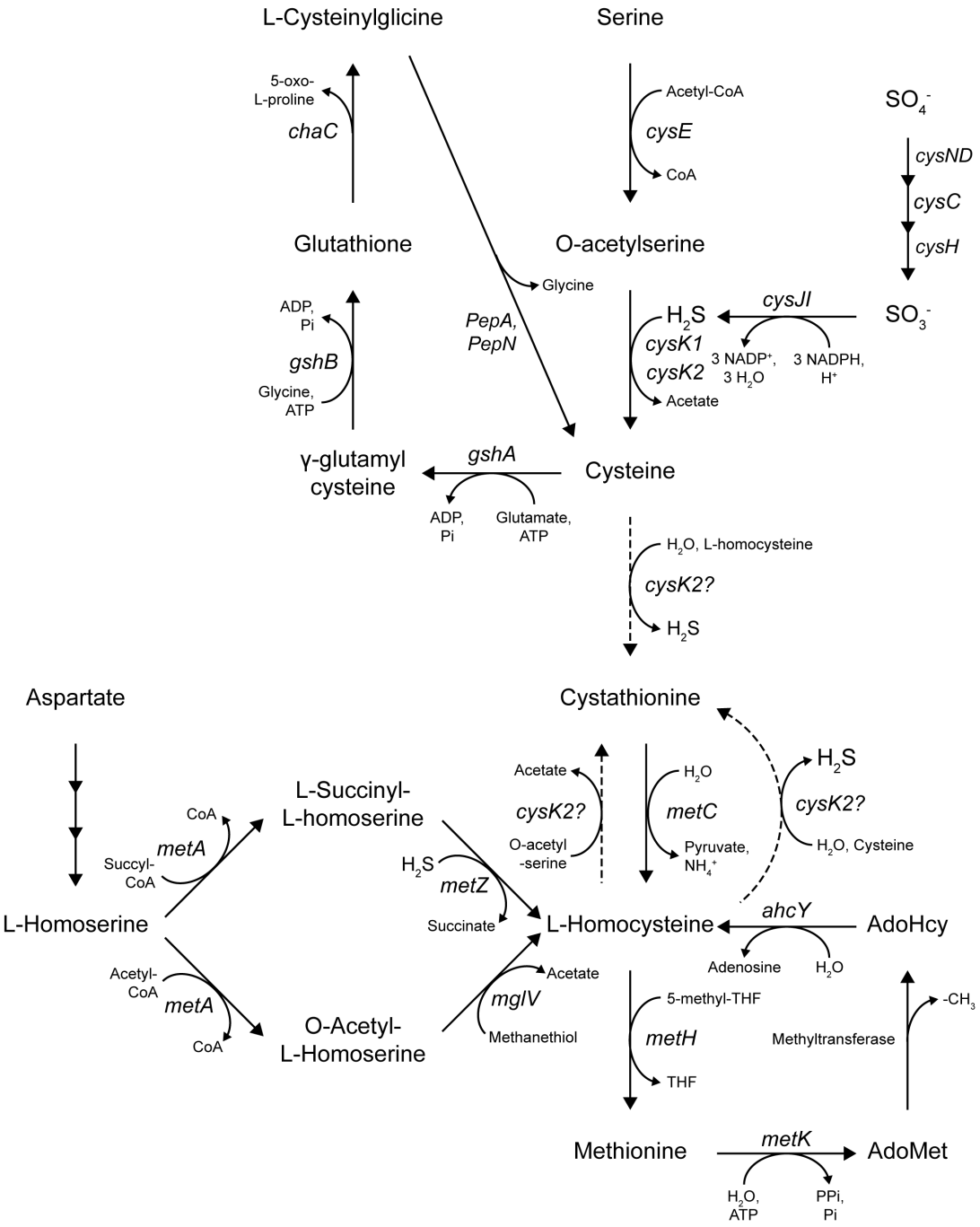


Figure 6.30 - Predicted cysteine metabolism in *B. ovis*

Illustration of the predicted molecules and enzymes involved in cysteine metabolism in *B. ovis*. Abbreviations: *achy*, adenosylhomocysteinase; *chaC,* glutathione-specific γ-glutamylcyclotransferase*;* CoA, coenzyme A*; cysC,* adenylylsulfate kinase; *cysE,* serine-O-acetyltransferase*; cysH,* 3'-phosphoadenlylysulfate reductase; *cysJI,* sulfate reductase*; cysK1/2,* cysteine synthase; *cysND*, sulfate adenylyltransferase; *gshA*, glutamate—cysteine ligase; *gshB*, glutathione synthetase; *metA*, homoserine transsuccinylase; *metC*, cystathionine β-lyase; *metH*, methionine synthase; *metK*, S-adenosylmethionine synthase; *metZ*, O-succinylhomoserine

Continuation: Figure 6.30 – Predicted cysteine metabolism in ovis sulfhydrylase; *mlgV*, O-acetylhomoserine (thiol)-lyase*; pepA*, aminopeptidase A; *pepN,* aminopeptidase N.

		I	II	III	IV	V
Conserved	OCBS	TAGNTG	GSGGT	TEGIGME	SSSG	RYLSK
	CBS	TSGNTG	GTGGT	VEGIGYD	GSSG	RYLST
	OASS	TSGNTG	GTGGT	IQGIGAG	ISSG	NYMTK
B. ovis	CysK1	TSGNTG	GTGGT	IQGLGAG	ISSG	RYLST
	CysK2	TAGNTG	GSGGT	AEGIGQG	GSSG	RYQSK
H.p.	OCBS (hp)	TAGNTG	GSGGT	IEGIGVE	SSSG	RYLSK

Figure 6.31 - Similarities of conserved regions of OCBS and OASS to *B. ovis* CysK1/2, and effect of cystathionine supplementation
Conserved regions identified by Devi and colleagues (257) in cysteine synthase (i.e. O-acetylserine sulfhydrylase, OASS), cystathionine b-synthase (CBS), and O-acetylserine-dependent CBS (OCBS) compared to CysK1 and CysK2 from *B. ovis*. Significantly different amino acids are underlined.

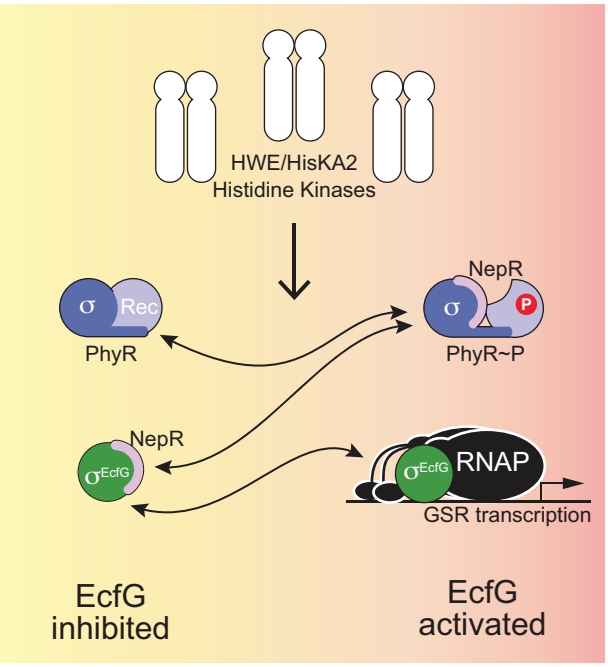


Figure 6.32 - Core regulators of the Alphaproteobacterial general stress response
In *Alphaproteobacteria*, the general stress response (GSR) is often controlled by an ensemble of HWE/HisKA_2 family sensor histidine kinases that modulate the phosphorylation state of PhyR. PhyR phosphorylation promotes a protein partner switch that sequesters the anti-sigma factor, NepR, thereby releasing the extracytoplasmic function (ECF) sigma factor, EcfG, to initiate

Continuation: Figure 6.32 – Core regulatoris of the Alphaproteobacterial general stress response
transcription of genes in the general stress response regulon. The core protein partner switch genes, *phyR*, *nepR* and *ecfG* are broadly conserved in *Alphaproteobacteria*. The number of kinases controlling the pathway is variable.

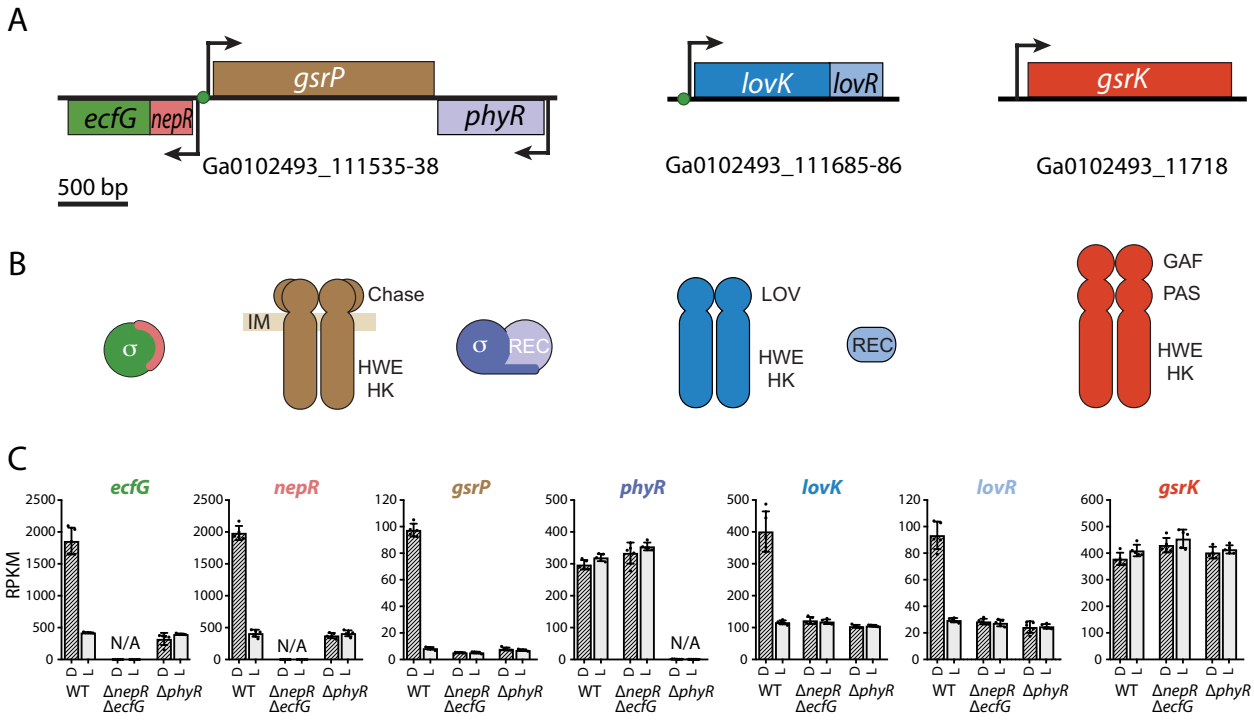


Figure 6.33 - GSR regulators and their expression in *E. litoralis* DSM 8509
A) Genes encoding the GSR regulators in *E. litoralis*. The core partner switch genes (*phyR*, *nepR*, and *ecfG*) together with a HWE kinase gene (*gsrP*) constitute the GSR locus (left). The *lovKR* operon and *gsrK* are not adjacent to each other or to the GSR locus on the chromosome. Gene locus numbers (GenBank accession CP017057) are listed below each gene diagram. Predicted EcfG-binding motifs are marked with green circles at the base of the transcription start arrows. **B)** Domain structure of the proteins encoded by the genes in (**A**). Colors match the gene colors. GsrP is predicted to encode a transmembrane histidine kinase and is depicted in the inner membrane (IM). All other proteins lack transmembrane domains and are predicted to be cytoplasmic. **C)** RNA-seq analyses of transcript abundance from the genes in (**A**). Reads per kilobase per million reads (RPKM) of each gene is plotted for wild-type (WT), Δ*nepR-ecfG* or Δ*phyR* strains grown in the dark (D - hashed bars) or the light (L – light grey bars). Mean ± s.d. is plotted. n=5 per treatment; individual values from experimental replicates are dots.

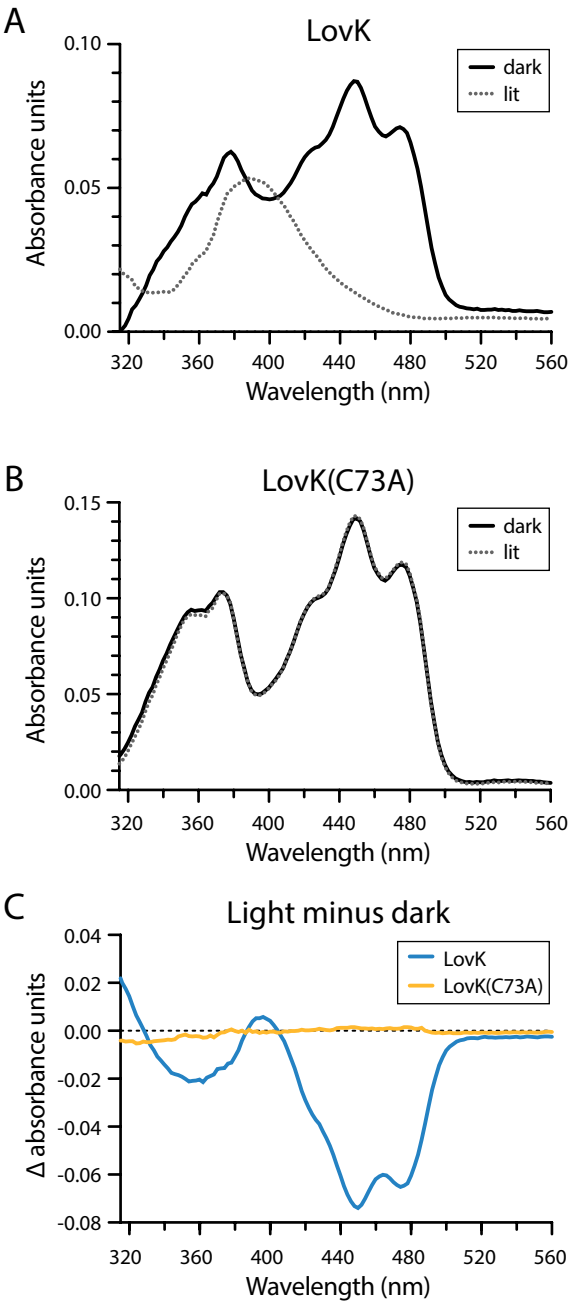


Figure 6.34 - Absorption spectroscopy provides evidence that LovK protein can function as a photosensor

A-B) Visible absorption spectrum of purified *E. litoralis* DSM 8509 (**A**) LovK and (**B**) LovK(C73A) proteins. Lit spectrum was collected immediately after illuminating the protein with blue light. The dark spectrum is the same protein that was not illuminated. **C)** Light minus dark difference absorption spectrum for LovK (blue line) and LovK(C73A) (yellow line). Photoexcitation of the wild-type protein results in a signature loss of absorbance between at 440 and 480 nm and an increase in absorbance at 396 (310, 313) due to cysteinyl-flavin C4(a) adduct formation (311, 312). The LovK(C73A) mutant protein lacks the conserved cysteine required for adduct formation and does not exhibit light induced changes in absorbance.

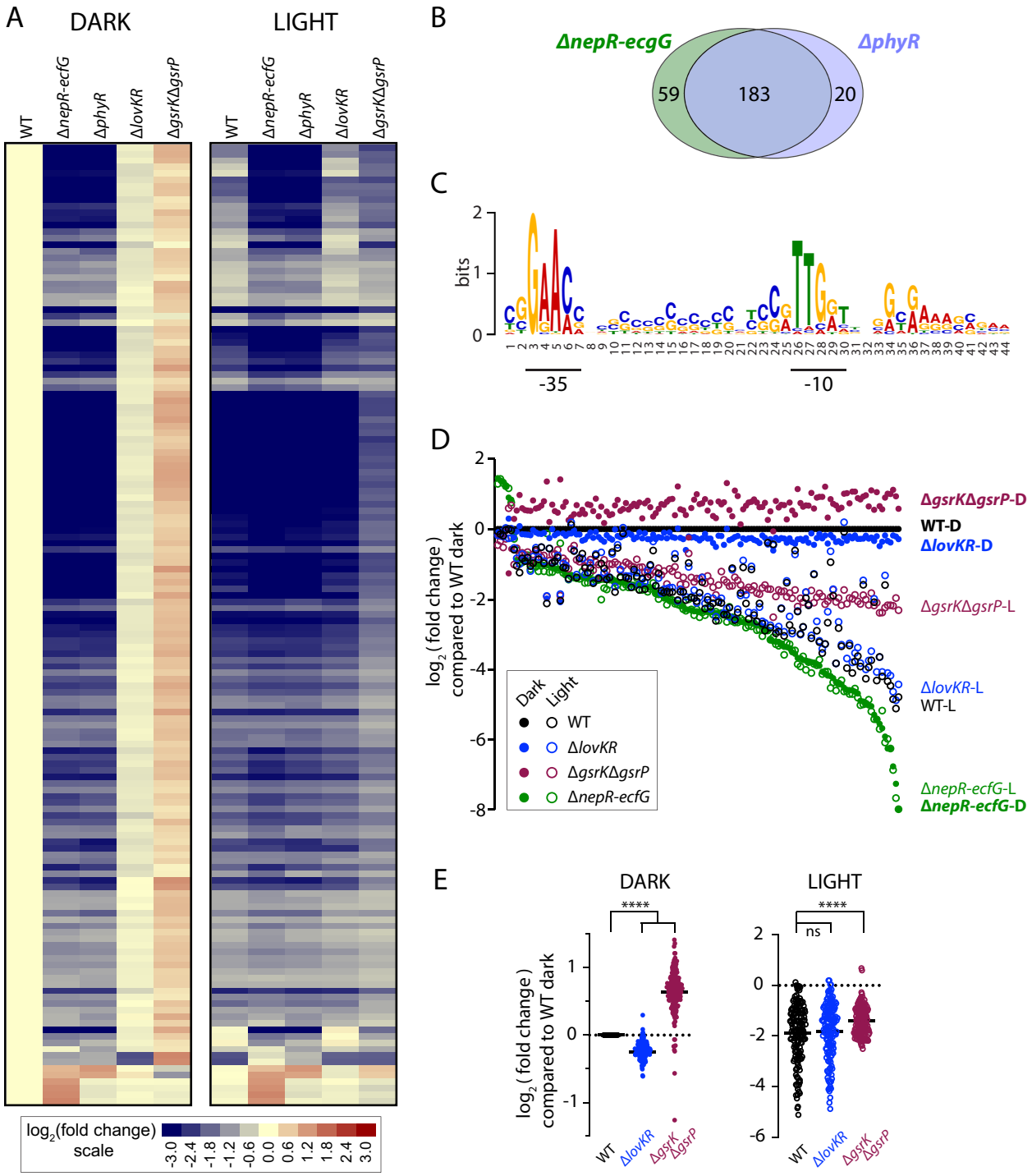


Figure 6.35 - General stress response regulon of *E. litoralis* DSM 8509

A) Heatmap showing relative expression of genes in the GSR regulon, defined by RNA-sequencing. Each row represents a gene, and each column represents a genotype-condition pair. Wild-type (WT) and mutant strains grown in the dark are on the left and strains grown under constant white light illumination (~60 µmol m^{-2} s^{-1}) are on the right. The color or each block represents the log_2(fold change) for an individual gene, where fold change is the ratio of $RPKM_{(indicated\ condition)}$ / $RPKM_{(WT\ dark)}$. Color scale is below the heat map. The heat map contains a

Continuation: Figure 6.35 – General stress response regulation of *E. litoralis* DSM 8509
subset of 148 genes in **Table S4** with a fold change > 2 in the Δ*nepR-ecfG* or the Δ*phyR* data sets, and a max group mean RPKM > 10. GSR regulatory genes that were deleted in these strains were excluded from the gene set presented in the heatmap. **B)** Venn diagram of the differentially regulated genes in the Δ*nepR-ecfG* or Δ*phyR* strains compared to wild type. The criteria for genes deemed to be differentially regulated was a fold change > 1.5 with a false-discovery rate (FDR) p-value < 0.01. These genes in are listed in **Table S4**. **C)** Logo of the ECF σ-type motif identified in the promoter region of a subset of genes in the GSR regulon; the logo was generated using MEME motif discovery tools. The -35 and -10 boxes (underlined) of this *E. litoralis* motif are consistent with previously described EcfG motifs (40, 42, 314). **D)** Relative expression of the 148 genes in panel A. Each position on the x-axis represents a gene ranked (left to right) by $\log_2$(fold change) in Δ*nepR-ecfG* dark relative to WT dark. Color-coded dots at each position represent the relative expression of a gene in different strain-condition combinations compared to WT dark (see key). Relative expression in dark-grown cultures are closed circles; light-grown cultures are open circles. **E)** Relative expression values for genes in the GSR regulon in WT, Δ*lovKR* and Δ*gsrK*Δ*gsrP* strains (grown in light or in dark) presented in a single column. Expression of each gene in these strain backgrounds is plotted relative to WT dark. The mean difference between each column was statistically assessed by one-way ANOVA followed by Dunnett's multiple comparison test; **** indicates $p<0.0001$. The key in **D** corresponds to both graphs in **E**.

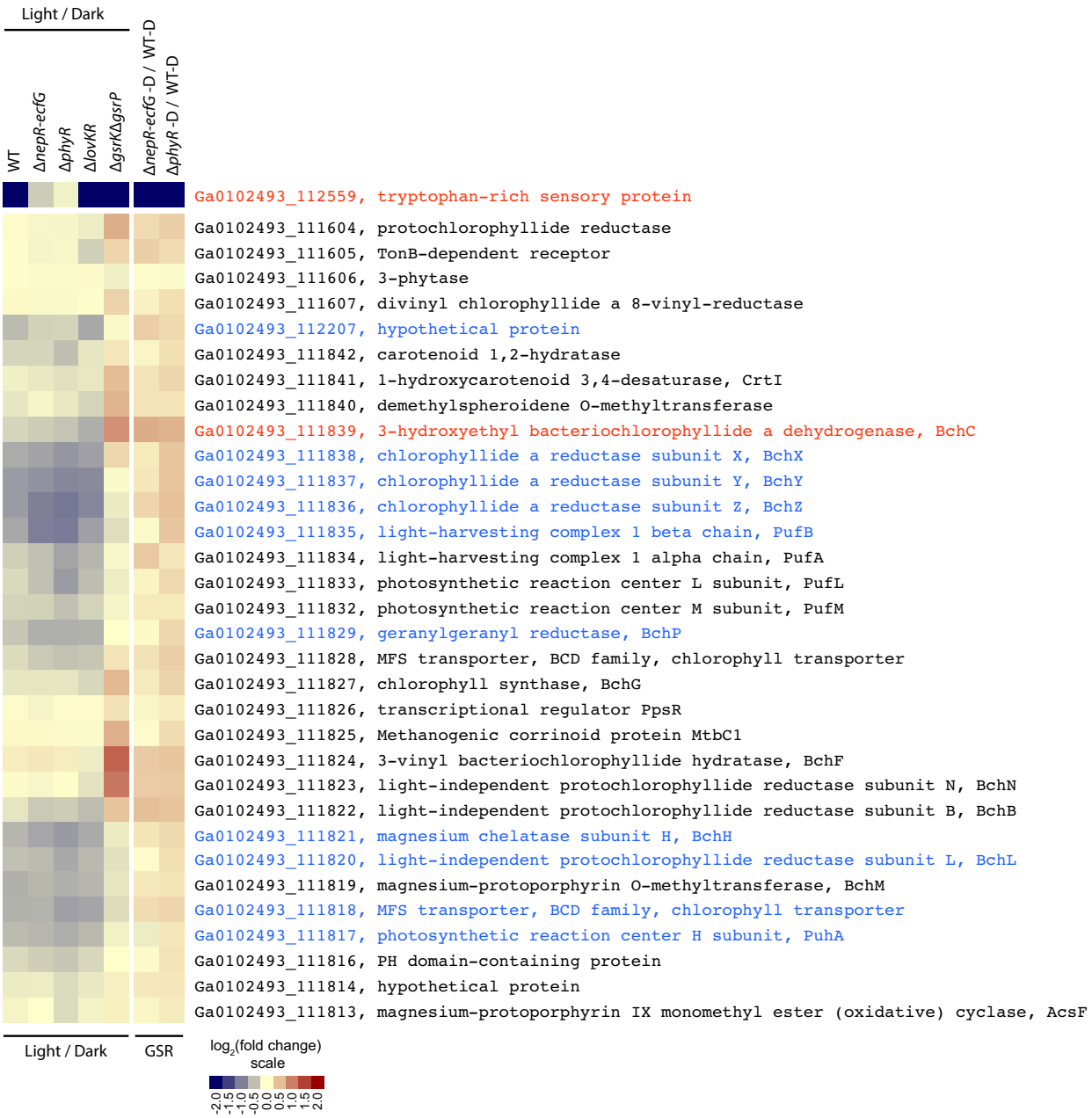


Figure 6.36 - Relative expression of genes involved in phototropy
Heatmap represents the same comparisons and same color scaling as in **Figure 6.37**. Genes were selected based on annotated functions in bacteriochlorophyll synthesis, phototrophy, or proximity to such genes. Genes are ordered by position on the chromosome. Genes in red text meet the cutoff criteria for inclusion in the GSR regulon (**Table S4**). Genes in blue text meet the cutoff criteria for inclusion in the light-dark regulon (**Table S5**).

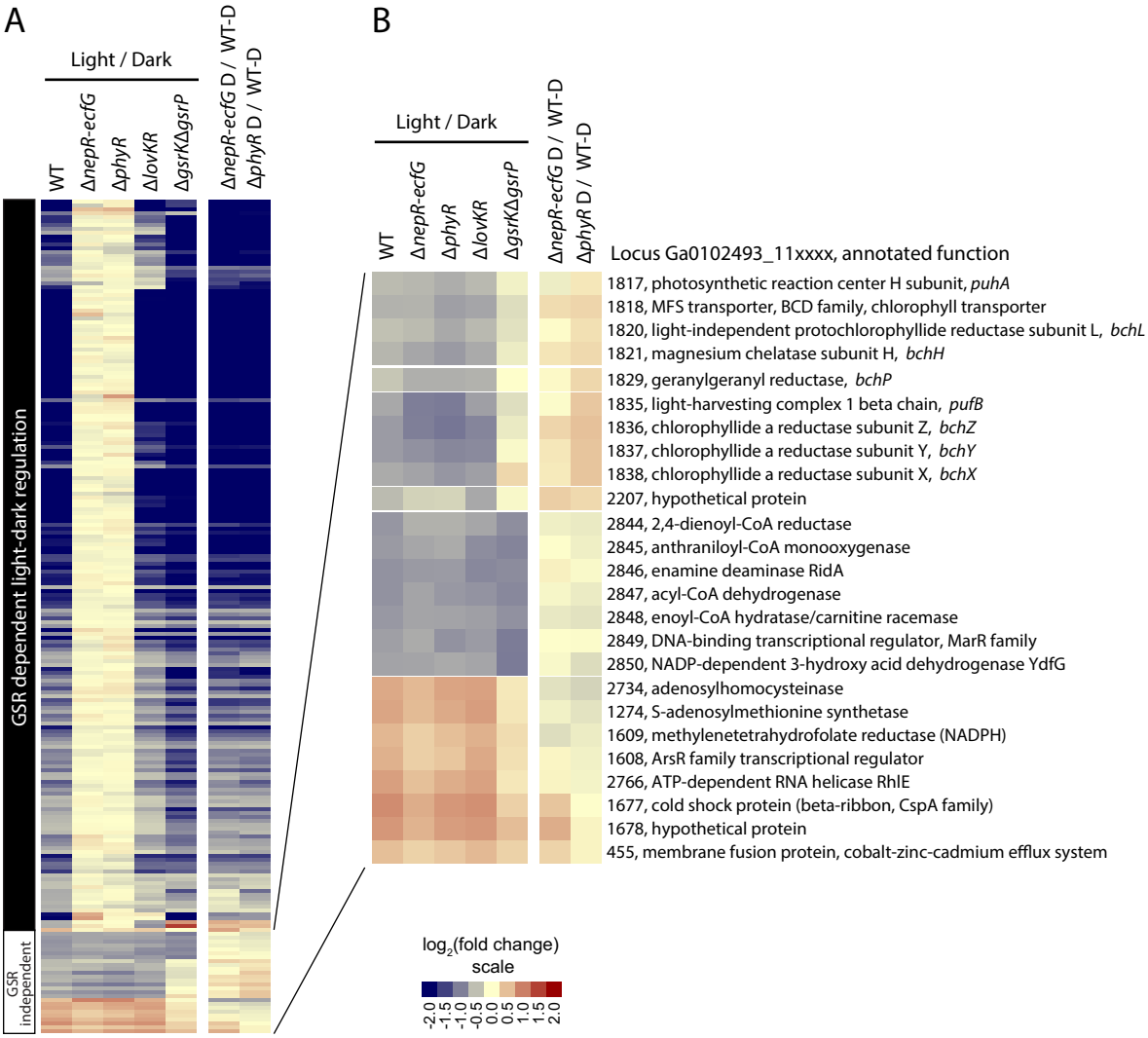


Figure 6.37 - The *E. litoralis* light-dark regulon overlaps with the GSR regulon
A) Heatmap of differentially expressed genes in wild-type (WT) *E. litoralis* DSM 8509 grown in constant ~60 µmol m^{-2} s^{-1} white light or in the dark (**D**). Cutoff criteria for differential expression are fold change > 1.4 and FDR p-value<0.01 (see **Table S5** for list of genes). Heatmap includes only genes with max mean RPKM >10. For each gene (in rows), the $\log_2$ ($RPKM_{light}$ / $RPKM_{dark}$) is shown for each genotype (columns). In addition, relative change in the Δ*nepR-ecfG* and Δ*phyR* strains compared to wild type (all grown in the dark) is presented to highlight congruence between light-regulated genes and the GSR regulon. **B)** Heatmap highlighting genes at the bottom of the cluster in panel **A**. These genes exhibit light-dependent regulation, but are not in the GSR regulon. Color scale corresponds to both panels.

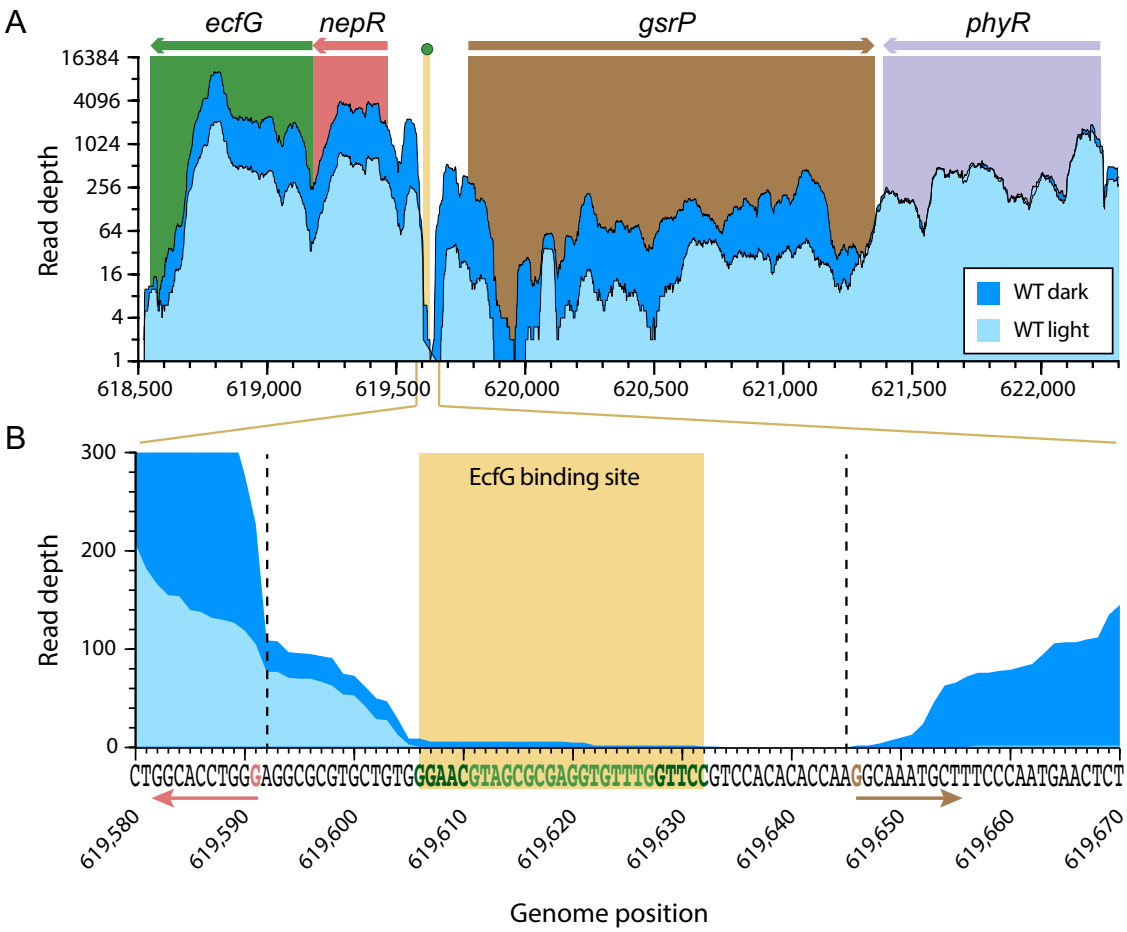


Figure 6.38 - Palindromic σ^{EcfG} binding site lies between *nepR-ecfG* and *gsrP*
A) RNA-seq reads mapped to the GSR locus from wild-type cells grown in the dark or in the light. Number of mapped reads are plotted as a function of genome position (GenBank accession CP017057). The genes in this region are colored as in **Figure 6.33.** The single σ^{EcfG} motif in this region is marked by a green dot and beige bar. Mapped read depth for *phyR* is similar in light and dark conditions, while RNA-seq reads are more abundant in the dark for *nepR-ecfG* and *gsrP*. **B)** Expansion of the intragenic region between *nepR-ecfG* and *gsrP*. Palindromic bases of the -35 and -10 sites of EcfG motif are dark green, and the intervening bases are lighter green. The beginning of the peak of reads for each transcript is marked with a vertical dashed line. The predicted +1 sites are colored and marked with an arrow to indicate the direction of transcription (colors correspond to the genes as in panel **A**).

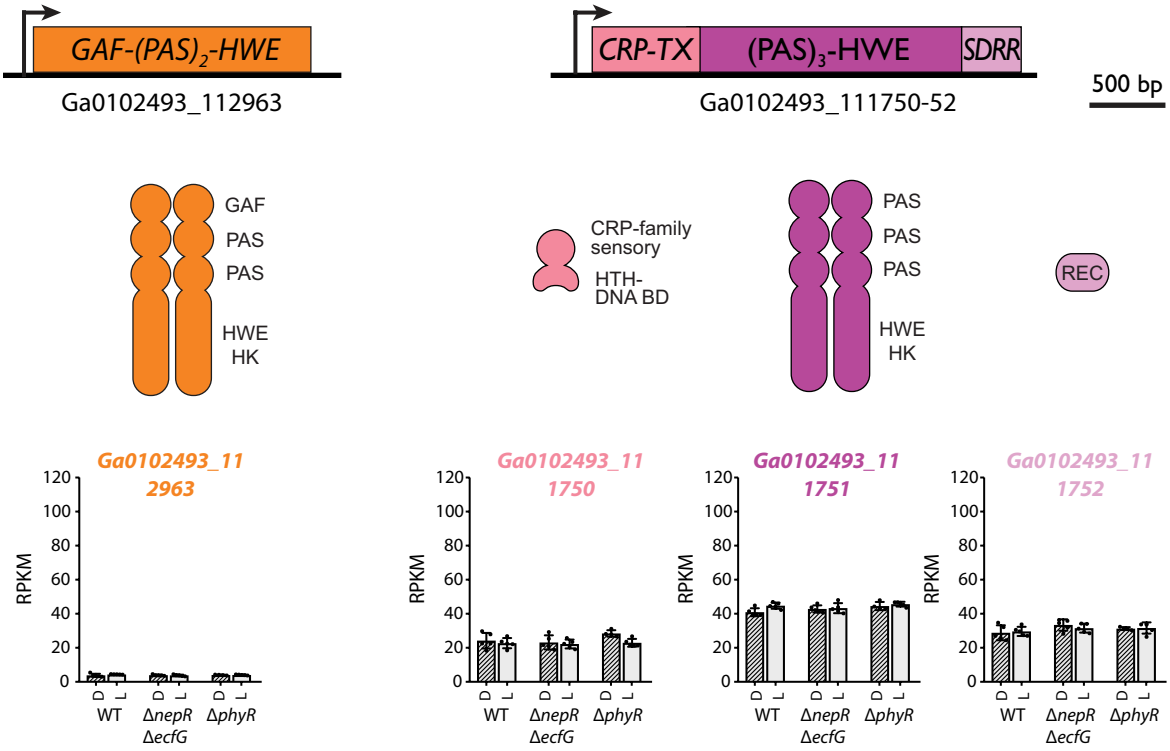


Figure 6.39 - Gene structure, protein domain structure and RNA-seq expression values for two additional unnamed HWE kinases encoded in the DSM 8509 genome
Features are drawn as in **Figure 6.33**.

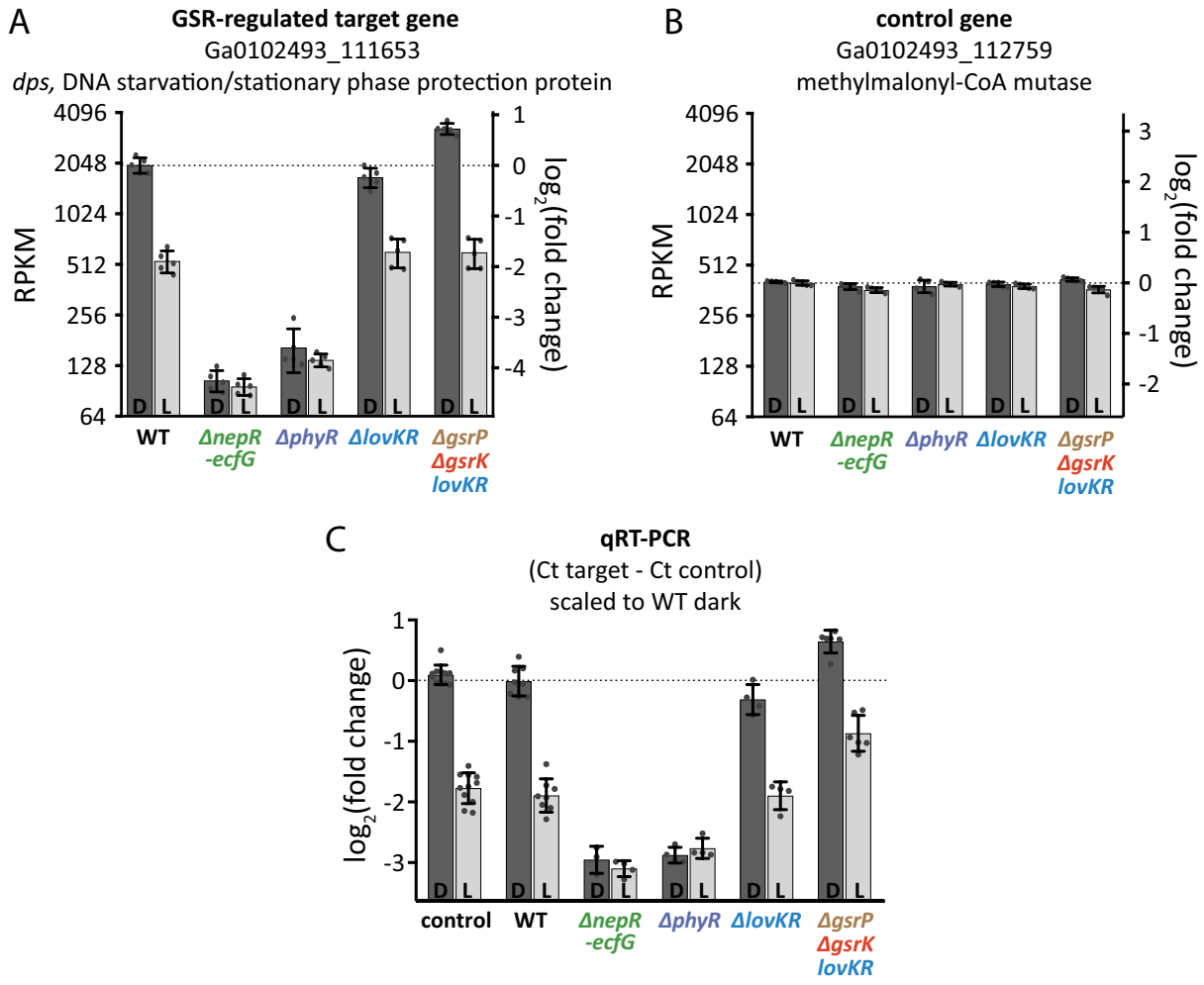


Figure 6.40 - Target and control genes used for qRT-PCR analysis of GSR transcription
A) RPKM values for *dps* (a GSR-regulated gene), extracted from RNA-seq experiments (**Table S3**) plotted on a $\log_2$ scale. Bars represent mean ± s.d. of 5 independent samples (dots). Log_2 (fold change) relative to wild type (WT) dark is scaled on the right y-axis. **B)** RPKM values for the methylmalonyl-CoA mutase gene used as the endogenous control gene for normalization plotted as in (**A**). **C)** qRT-PCR analysis of the same genotype-condition combinations assayed by RNA-seq in (**A**) and (**B**). These data are extracted from the same set of experiments presented in **Figure 6.38** and are presented here for direct comparison to panel (A). As in **Figure 6.38**, each measurement represents the $(Ct_{dps} - Ct_{control})_{sample}$ – average $(Ct_{dps} - Ct_{control})_{WT\text{-}D}$. This results in a $\log_2$(fold change) compared to WT-dark. Strains were grown and assayed together. Each sample was assayed in triplicate. Points represent the average value for each sample. Bars represent mean ± s.d. of the samples in each condition. As an internal control, the same wild type-dark and wild type-light samples were assayed on every plate (control).

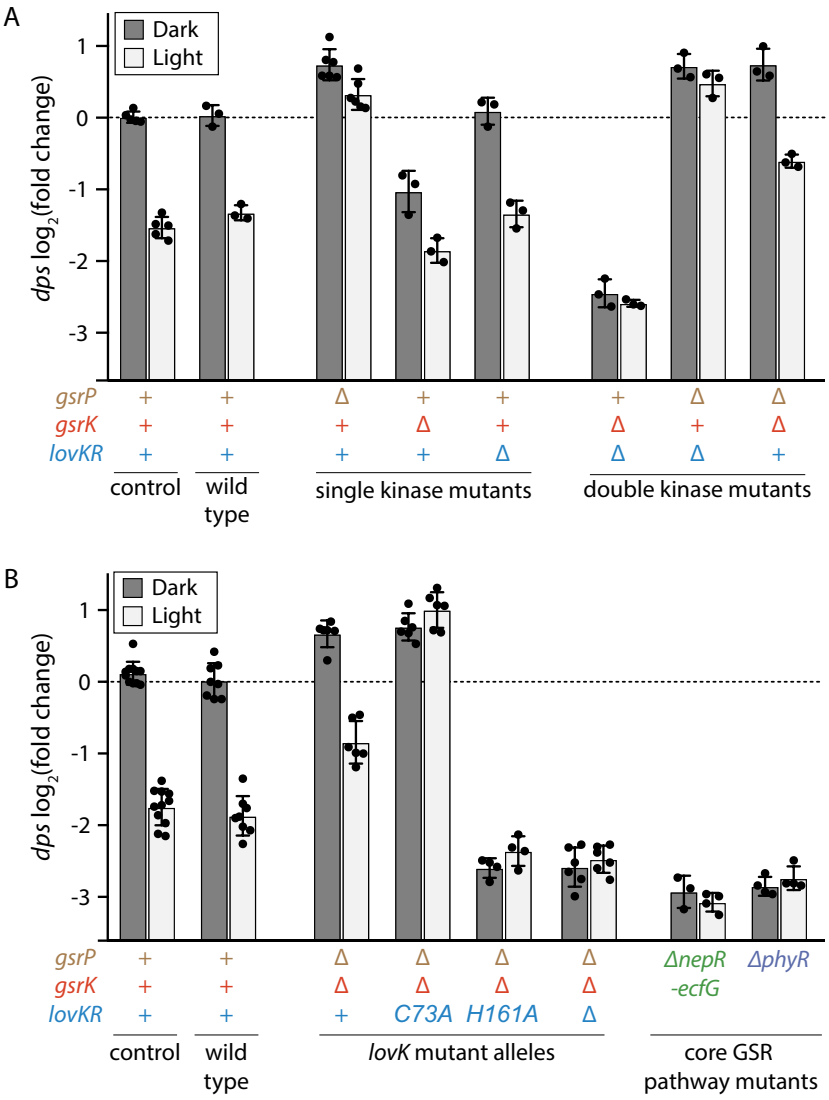


Figure 6.41 - Combinatorial control of GSR transcription by three HWE-family sensor histidine kinases

A-B) qRT-PCR quantification of the levels of, *dps*, a GSR-regulated transcript. *dps* transcript measurements were carried out on RNA isolated from (**A**) a set of single and double HWE kinase mutant strains and (**B**) *lovK* mutant, Δ*nepR-ecfG* and Δ*phyR* strains grown in the dark (D, dark grey bars) or the light (L, light grey bars). Each measurement represents the $(Ct_{dps} - Ct_{control})_{sample}$ – average $(Ct_{dps} - Ct_{control})_{WT\text{-}D}$, which yields a log_2(fold change) in *dps* level relative to WT dark. Strains in each panel were grown and assayed together. Each sample was assayed in triplicate. Points represent the average value for each sample. Bars represent mean ± s.d. of all samples in each condition. As an internal control, a pair of wild type-dark and wild type-light samples was assayed on every plate (control).

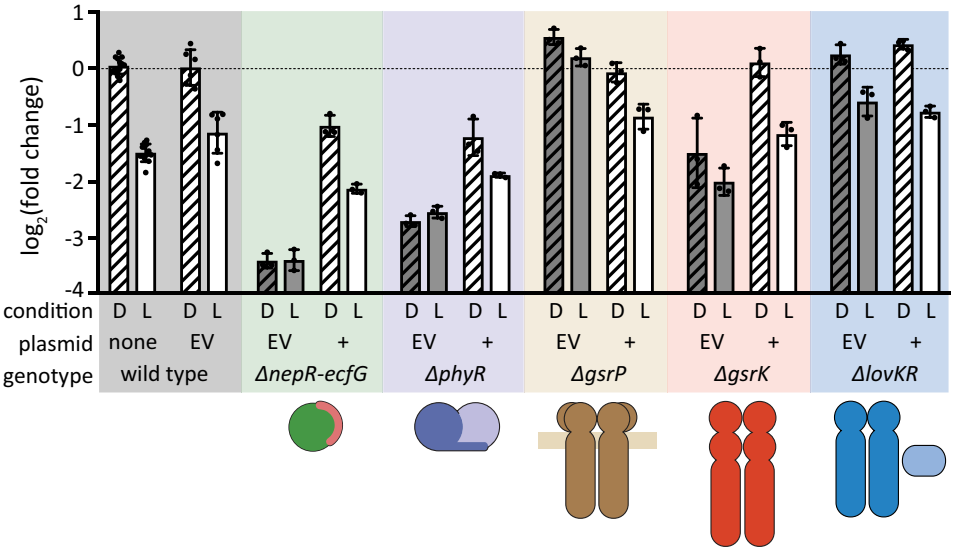


Figure 6.42 - Complementation of GSR transcription defects in strains with single deletions of GSR regulators

qRT-PCR analysis of *dps* expression as a measurement of GSR transcription. All strains, including wild type (WT), carry a replicating plasmid and were grown in the presence of gentamycin to select for plasmid maintenance. Plasmids either carried the deleted gene under the control of its endogenous promoter (+) or were empty vectors (EV) to control for plasmid and selection effects. Grey shaded bars highlight mutant strains carrying the empty vector. The chromosomal genotype of the strains in each colored block is indicated at the bottom. All strains were grown in the dark (D – striped bars) or in the light (L – open bars). Data are presented as in **Figure 6.38 and Figure 6.40C.** The internal control RNA samples assayed on every plate in this experiment were from wild type samples that did not carry a plasmid (far left bars).

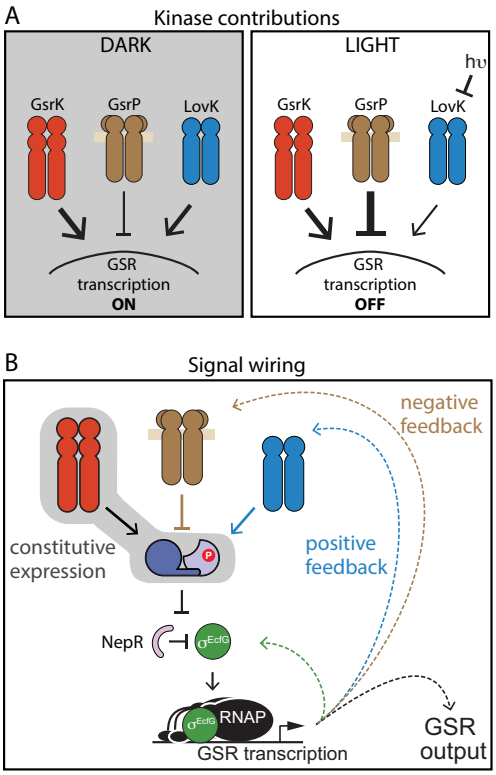


Figure 6.43 - Model of GSR regulation in E. litoralis DSM 8509
A) Three HWE-family sensor histidine kinases (LovK, GsrK, and GsrP) control output of σ^{EcfG}-dependent transcription. The weight of arrows shows the relative regulatory contribution of each sensor kinase under light and dark conditions. **B)** Model of GSR control network including protein interactions (solid lines) and transcriptional output (dashed lines). Most, but not all, of the GSR pathway regulators are subject to feedback regulation. PhyR, which serves to integrate signals, and GsrK, the primary activating kinase, are not under transcriptional control of σ^{EcfG}, and are instead constitutively expressed. This feature of the GSR system in *E. litoralis* DSM 8509 likely necessitates a negative regulator to prevent constitutive activation of GSR transcription; GsrP fulfills this function.

7. APPENDIX C – TABLES

Table 7.1 - Growth of independent spontaneous mutants derived from forward genetic selection grown in an air incubator (0.04% CO_2) after isolation[a]

Exp.[b]	Samples	OD_{600} at 24hrs	OD_{600} at 48hrs	WGS Samples[c]
Bov	WT1	0	0.001	Parent
	WT2	0	0	
Plate	1	0.011	1.645	Mutant 01
liq1	2	0.026	1.558	Mutant 02
liq2	3	0.01	1.654	Mutant 03
	4	0.015	1.664	
	5	0.014	1.654	
	6	0.022	1.655	Mutant 14
	7	0.007	1.64	
	8	0.014	1.657	Mutant 05
	9	0.005	1.617	Mutant 04
	10	0.009	1.646	
	11	0.01	1.647	
	12	0.012	1.655	Mutant 06
liq3	13	0.044	1.633	Mutant 07
	14	0.02	1.639	Mutant 08
	15	0.018	1.627	
	16	0.027	1.628	
	17	0.012	1.617	Mutant 09
	18	0.015	1.609	
	19	0.18	1.642	Mutant 15
	20	0.012	1.629	
	21	0.26	1.606	Mutant 10
liq4	22	0.021	1.651	Mutant 11
	23	0.024	1.593	
	24	0.026	1.554	Mutant 12
	25	0.038	1.635	Mutant 13
	26	0.027	1.635	
	27	0.028	1.635	Mutant 16
	28	0.068	1.638	
	29	0.028	1.634	

[a]Cultures were inoculated at an OD_{600} of 1.5×10^{-5}

[b]Spontaneous independent mutants from independent experiments (Exp.). "Plate" refers to selection conducted on solid media instead of liquid, by transferring plated cells from 5% to 0.04% CO_2 incubators. Liquid (liq) 1-4 are four independent experiments were cells were inoculated in broth. See **Results** and **Materials and methods**. "Bov" is the parent strain (*Brucella ovis* ATCC 25840) where two independent tubes were inoculated for comparison with the mutants.

[c]Samples indicated were sent for whole genome sequencing (WGS).

Table 7.2 - Sequence polymorphisms[a] between 16 independent *B. ovis* mutants that grow without CO_2 supplementation and the wild-type parent strain[b]

Mut #	Chr.	Pos.	Mutation	Locus[c]	Annot.	Desc.
01	I	805,096	C → T	BOV_RS03980	NADH-quinone oxidoreductase subunit B	T188T (ACC → ACT)
	I	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
02[d]	I	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
03	I	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
04	I	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
05	I	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
06	I	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
07	I	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
08	I	1,768,984	-A	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA3*$_{BOV}$ (frameshift)
09	I	301,445	A → G	BOV_RS0175	IS5/IS1182 family transposase	L224P (CTT → CCT)
	I	301,958	A → C	BOV_RS01475	sensor histidine kinase	V53G (GTT → GGT)
	I	420,434	G → T	BOV_RS02035 ← / ← BOV_RS16215	hypothetical protein / hypothetical protein	intergenic
	I	1,725,552	C → T	BOV_RS08460	branched-chain amino acid ABC transporter ATP-binding protein	G349D (GGC → GAC)

Continuation: Table 7.2 - Sequence polymorphisms[a] between 16 independent *B. ovis* mutants that grow without CO_2 supplementation and the wild-type parent strain[b]

	I	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
	II	117,883	(T)8 → (T)7	BOV_RS10890 → / → BOV_RS12585	HTH-type quorum sensing-dependent transcriptional regulator VjbR	intergenic
	II	676,505	G → A	BOV_RS13635	sulfurtransferase FdhD	pseudogene (415/810 nt)
	II	908,237	C → T	BOV_RS14715 → / → BOV_RS14720	α-hydroxy-acid oxidizing protein lldD / porin family protein	intergenic
10	I	805,096	C → T	BOV_RS03980	NADH-quinone oxidoreductase subunit B	T188T (ACC → ACT)
	I	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
11	I	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)
12[d]	I	1,712,213	G → T	BOV_RS08400	polyprenyl synthetase family protein	A16S (GCC → TCC)
	I	1,768,979	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA4*$_{BOV}$ (frameshift)
13	I	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
14	I	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
15	I	1,768,980	-G	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA2*$_{BOV}$ (frameshift)
16	I	1,768,982	-C	BOV_RS08635	β-carbonic anhydrase - *bcaA*$_{BOV}$	*bcaA*$_{BOV}$ → *bcaA1*$_{BOV}$ (frameshift)

Continuation: Table 7.2 - Sequence polymorphisms[a] between 16 independent *B. ovis* mutants that grow without CO_2 supplementation and the wild-type parent strain[b]

Abbreviations: Mut# = mutant number, Chr .= Chromosome, Pos. = Position, Annot. = Gene Annotation, Descr. = Description.

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

[b] We observed four differences between the *B. ovis* ATCC 25840 sequence deposited in GenBank and the sequence of our ATCC 25840 parent strain: In 100% of the reads, there was an insertion of a single G in BOV_RS01360 on chromosome I (NC_009505), at base 282,040; an insertion of a single C in an intergenic region on chromosome II (NC_009504), at position 20; and an insertion of a single G in an intergenic region at base 459,010, also on chromosome II. These differences were also present in all derived 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 base 118,034 of chromosome II. Each derived mutant strain carried one of several different *vjbR* polymorphisms, each present in 100% of the sequencing reads.

[c]For intergenic polymorphisms, ← 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.

Table 7.3 - *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 Tn 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]Chao2 estimate of the total number of barcodes in the library

[b]Sites on opposite strand count as unique

[c]Genes > 300nt without good insertions were counted as 'unhit'

Table 7.4 - *B. abortus* biovar association with *bcaA* allele clusters

Biovar	CO_2 dep.[a]	Tot.	Cl. 1 (*bcaA*$_{BAB}$)	Cl. 2	Cl. 3	Cl. 4	Cl. 5	Cl. 6	Cl. 7[b] (*bcaA*$_{BAB2308}$)	Cl. 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
N/A		105[b] (28%)	57 (55%)	23 (22%)	0	11 (10%)	4 (4%)	2 (2%)	5 (5%)	2 (2%)

Abbreviations: Cl. = Clutser, Tot. = Total, CO_2 dep. = CO_2 dependence.
[a]Symbols indicating CO_2 dependence as per Alton *et al.*(215): + = CO_2 dependent; + (–) = mostly CO_2 dependent; – = CO_2 independent.
[b]Two strains are annotated as 2308, thus are possibly biovar 1 and redundant. Biovar was not assigned.
[c]One strain is annotated as “atypical”

Table 7.5 - Gene and pseudogene comparison across *B. ovis* and *B. abortus* ATCC 2308 strains

Group	Gene	*B. ovis* old locus	*B. ovis* new locus	*B. abortus* locus	Gene length[a]	Polymorphisms between *B. ovis* and *B. abortus*[b]				Polymorphisms within *B. ovis* genomes	
						SNPs		Single nt Indels	Indels >1nt		
						Tot	Syn[c]			# site(s)[d]	# genomes[e]
Pseudogenes	*ureF2*	BOV_1316	BOV_RS06515	BAB_RS22515	732	3	/	2	0	0	0
	ureT	BOV_1319	BOV_RS06530	BAB_RS22530	1050	2	/	0	1 (56nt)	1	15
	ureG1	BOV_0287	BOV_RS01465	BAB_RS17375	627	1	/	0	0	0	0
	ureE2	BOV_1315	BOV_RS06510	BAB_RS22510	606	2	/	2	0	0	0
	ureC1	BOV_0284	BOV_RS01450	BAB_RS17360	1713	5	/	0	1 (30nt)	0	0
	pckA	BOV_2009	BOV_RS09880	BAB_RS25895	1476	8	/	1	0	1	1
	eryA	BOV_A0811	BOV_RS14445	BAB_RS28140	1554	6	/	0	0	0	0
	eryD	BOV_A0814	BOV_RS14460	BAB_RS28125	951	4	/	0	1 (7nt)	0	0
	eryF	BOV_A0805	BOV_RS14425	BAB_RS28160	948	4	/	0	1 (9nt)	0	0
	eryG	BOV_A0806	BOV_RS14430	BAB_RS28155	1041	1	/	0	1 (2nt)	0	0
	gluP	BOV_A0172	BOV_RS11215	BAB_RS27245	1239	3	/	1	0	0	0
	ccoO	BOV_0378	BOV_RS01915	BAB_RS17800	732	1	/	1	0	1	11
	coxB	/	BOV_RS02370	BAB_RS18265	889	2	/	1	0	0	0
	coxG	BOV_0478	BOV_RS16275	BAB_RS18290	879	2	/	1	2 (56nt, 6nt)	0	0
	ctaE	/	BOV_RS11490	BAB_RS30835	570	2	/	0	2 (12nt, 27nt)	0	0
	ctaG	BOV_0477	BOV_RS02390	BAB_RS18280	606	1	/	0	1 (5nt)	0	0
	norB	/	BOV_RS11505	BAB_RS30820	1350	6	/	0	1 (80nt)	1	1

Continuation: Table 7.5 - Gene and pseudogene comparison across *B. ovis* and *B. abortus* ATCC 2308 strains

	copA/ fixI	/	BOV_RS01885	BAB_RS17775	2259	5	/	0	1 (153nt)	0	0
	bcaA	/	BOV_RS08635	BAB_RS24650	645	0	/	1	1 (2nt)	0	0
Urease genes	*ureD1*	BOV_0281	BOV_RS01430	BAB_RS17340	843	2	1	0	0	0	0
	ureA1	BOV_0202	BOV_RS01435	BAB_RS17345	303	1	0	0	0	0	0
	ureB1	BOV_0283	BOV_RS01445	BAB_RS17355	306	0	0	0	0	0	0
	ureE1	BOV_0285	BOV_RS01455	BAB_RS17365	516	5	3	0	0	0	0
	ureF1	BOV_0286	BOV_RS01460	BAB_RS17370	688	2	0	0	0	1	1
	ureA2	BOV_1312	BOV_RS06495	BAB_RS22495	303	1	0	0	0	0	0
	ureB2	BOV_1313	BOV_RS06500	BAB_RS22500	480	2	0	0	0	0	0
	ureC2	BOV_1314	BOV_RS06505	BAB_RS22505	1722	5	1	0	0	0	0
	ureG2	BOV_1381	BAB_RS22520	BAB_RS22520	639	1	1	0	0	0	0
	ureD2	BOV_1318	BOV_RS06525	BAB_RS22525	915	4	1	0	1 (6nt)	0	0
Type IV Secretion System	*virB11*	BOV_A0054	BOV_RS10610	BAB_RS26635	1086	2	2	0	0	0	0
	virB10	BOV_A0055	BOV_RS10615	BAB_RS26640	1167	3	1	0	1 (24nt)	1	1
	virB9	BOV_A0056	BOV_RS10620	BAB_RS26645	870	0	0	0	0	1	1
	virB8	BOV_A0057	BOV_RS10625	BAB_RS26650	720	2	1	0	0	0	0
	virB7	/	BOV_RS10630	BAB_RS26655	174	0	0	0	0	0	0
	virB6	BOV_A0058	BOV_RS10635	BAB_RS26660	1044	1	1	0	0	1	11
	virB5	BOV_A0059	BOV_RS10640	BAB_RS26665	717	4	2	0	0	0	0
	virB4	BOV_A0060	BOV_RS10645	BAB_RS26670	2496	4	1	0	0	2	2[f]
	virB3	BOV_A0061	BOV_RS10650	BAB_RS26675	351	1	1	1	0	0	0

Continuation: Table 7.5 - Gene and pseudogene comparison across *B. ovis* and *B. abortus* ATCC 2308 strains

virB2	BOV_A0062	BOV_RS10655	BAB_RS26680	318	0	0	0	0	0	0
virB1	BOV_A0063	BOV_RS10660	BAB_RS26685	717	2	0	0	0	0	0

[a]In *B. abortus* ATCC 2308
[b]Number of polymorphisms between all 17 sequenced *B. ovis* clinical isolates (see **Data Sheet 1**) and *B. abortus* ATCC 2308 strain (GenBank accessions NC_007618 and NC_007624).
[c]Syn = synonymous mutations. In the case of pseudogenes, synonymous mutations were not noted.
[d]Number of polymorphic sites among sequences from *B. ovis* isolates
[e]Number of *B. ovis* sequences that harbor the polymorphism
[f]Two genomes have distinct polymorphisms

Table 7.6

Timepoint	**OD_{600}**	**Volume**	**Number of cells collected**	**Cells in each PCR rxn**
T0	0.6	100 ml	4×10^9	8×10^7
T1	0.05	1 ml	6.6×10^9	1.3×10^8
T2	0.12	1 ml	8×10^9	1.6×10^8
T3	0.9	0.5 ml	6×10^9	1.2×10^8
T4	2.4	0.5 ml	8×10^9	1.6×10^8

rxn = reaction;

Table 7.7 - Genome characteristics of select *Erythrobacter* spp. isolates

Isolate [a]	Source (reference)	Genome status	Genome length (Mb)	% GC	Phototrophy genes	LOV domain [b]	LOV kinases (SDRR – orphan) [c]	HWE / HisKA2 kinases [d]
E. longus DSM 6997 (Och 101)	Seaweed in Aburatsubo bay, Kanagawa, Japan (343)	WGS; 14 contigs (298)	3.6	57	YES	0	0	2
E. litoralis DSM 8509	Cyanobacterial mat, Supralitoral zone, Texal, Netherlands (301, 302)	WGS; 22 contigs (298) -- Complete (this reference)	3.21	65	YES	1	1-0	5
Erythrobacter sp. SD-21	Surface sediment, San Diego Bay, USA (344)	WGS; 19 contigs	2.97	63	NO	1	0-1	16
Erythrobacter sp. NAP1	Surface water, Northwest Atlantic Ocean (345)	WGS; 4 contigs (346)	3.26	61	YES	1	0-1	6
E. litoralis HTCC2594	10 m deep Sargasso Sea, Atlantic Ocean (297)	Complete (297)	3.05	63	NO	4	1-2	8

[a] Isolates sequenced at the time this work was initiated

[b] Extracted from Glantz *et al.*, 2016 (275)

[c] Subset of genes identified in Glantz *et al.*, 2016 (275) with kinase domains in an operon with a single domain response regulator (SDRR) or encoded as an orphan gene. All of these are HWE/HisKA_2-family kinases.

[d] Identified in the MIST database 2.0 (307)

8. REFERENCES

1. Kolter R, Siegele DA, Tormo A. 1993. The stationary phase of the bacterial life cycle. Annu Rev Microbiol 47:855–874.
2. Huisman G, Siegele DA, Zambrano MM, Kolter R. 1996. Morphological and Physiological Changes During Stationary Phase.
3. Finkel SE. 2006. Long-term survival during stationary phase: evolution and the GASP phenotype. Nat Rev Microbiol 4:113–120.
4. Lange R, Hengge-Aronis R. 1991. Growth phase-regulated expression of bolA and morphology of stationary-phase Escherichia coli cells are controlled by the novel sigma factor sigma S. J Bacteriol 173:4474–4481.
5. Jenkins DE, Schultz JE, Matin A. 1988. Starvation-induced cross protection against heat or H2O2 challenge in Escherichia coli. J Bacteriol 170:3910–3914.
6. Jenkins DE, Chaisson SA, Matin A. 1990. Starvation-induced cross protection against osmotic challenge in Escherichia coli. J Bacteriol 172:2779–2781.
7. Rinas U, Hellmutii K, Kang R, Seeger A, Schlieker H. 1995. Entry of Escherichia coli into stationary phase is indicated by endogenous and exogenous accumulation of nucleobases. Appl Environ Microbiol 61:4147–4151.
8. Lange R, Hengge-Aronis R. 1991. Identification of a central regulator of stationary-phase gene expression in Escherichia coli. Mol Microbiol 5:49–59.
9. Gentry DR, Hernandez VJ, Nguyen LH, Jensen DB, Cashel M. 1993. Synthesis of the stationary-phase sigma factor σ(s) is positively regulated by ppGpp. J Bacteriol 175:7982–7989.
10. Schellhorn HE. 2020. Function, Evolution, and Composition of the RpoS Regulon in Escherichia coli. Front Microbiol 11:560099.
11. Battesti A, Majdalani N, Gottesman S. 2011. The RpoS-mediated general stress response in Escherichia coli. Annu Rev Microbiol2011/06/07. 65:189–213.
12. Loewen PC, Hengge-Aronis R. 1994. The role of the sigma factor sigma S (KatF) in bacterial global regulation. Annu Rev Microbiol 48:53–80.
13. Gottesman S. 2019. Trouble is coming: Signaling pathways that regulate general stress responses in bacteria. J Biol Chem 294:11685–11700.
14. Galperin MY. 2018. What bacteria want. Environ Microbiol 20:4221–4229.
15. Franze de Fernandez MT, Hayward WS, August JT. 1972. Bacterial proteins required for replication of phage Q ribonucleic acid. Pruification and properties of host factor I, a ribonucleic acid-binding protein. J Biol Chem 247:824–831.
16. Muffler A, Fischer D, Hengge-Aronis R. 1996. The RNA-binding protein HF-I, known as a host factor for phage Qbeta RNA replication, is essential for rpoS translation in

Escherichia coli. Genes Dev 10:1143–1151.

17. Nogueira T, Springer M. 2000. Post-transcriptional control by global regulators of gene expression in bacteria. Curr Opin Microbiol 3:154–158.

18. Hengge-Aronis R. 1996. Back to log phase: σS as a global regulator in the osmotic control of gene expression in Escherichia coli. Mol Microbiol 21:887–893.

19. Milewska K, Krause K, Szalewska-Pałasz A. 2020. The stringent response of marine bacteria – assessment of (p)ppGpp accumulation upon stress conditions. J Appl Genet 61:123–130.

20. Pacios O, Blasco L, Bleriot I, Fernandez-Garcia L, Ambroa A, López M, Bou G, Cantón R, Garcia-Contreras R, Wood TK, Tomás M. 2020. (p)ppGpp and Its Role in Bacterial Persistence: New Challenges. Antimicrob Agents Chemother 64:e01283-20.

21. Okada Y, Makino S, Tobe T, Okada N, Yamazaki S. 2002. Cloning of rel from Listeria monocytogenes as an Osmotolerance Involvement Gene. Appl Environ Microbiol 68:1541 LP – 1547.

22. Cashel M, Gallant J. 1969. Two compounds implicated in the function of the RC gene of Escherichia coli. Nature 221:838–841.

23. 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 68:1128–1148.

24. Chatterji D, Ojha AK. 2001. Revisiting the stringent response, ppGpp and starvation signaling. Curr Opin Microbiol 4:160–165.

25. Dozot M, Boigegrain RA, Delrue RM, Hallez R, Ouahrani-Bettache S, Danese I, Letesson JJ, De Bolle X, Köhler S. 2006. The stringent response mediator Rsh is required for Brucella melitensis and Brucella suis virulence, and for expression of the type IV secretion system virB. Cell Microbiol 8:1791–1802.

26. Kim S, Watanabe K, Suzuki H, Watarai M. 2005. Roles of Brucella abortus SpoT in morphological differentiation and intramacrophagic replication. Microbiology 151:1607–1617.

27. Nazir A, Harinarayanan R. 2016. (p)ppGpp and the bacterial cell cycle. J Biosci 41:277–282.

28. Choy HE. 2000. The study of guanosine 5'-diphosphate 3'-diphosphate-mediated transcription regulation in vitro using a coupled transcription-translation system. J Biol Chem 275:6783–6789.

29. Durfee T, Hansen AM, Zhi H, Blattner FR, Ding JJ. 2008. Transcription profiling of the stringent response in Escherichia coli. J Bacteriol 190:1084–1096.

30. 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.

31. Lesley JA, Shapiro L. 2008. SpoT regulates DnaA stability and initiation of DNA replication in carbon-starved Caulobacter crescentus. J Bacteriol 190:6867–6880.

32. Chiaverotti TA, Parker G, Gallant J, Agabian N. 1981. Conditions that trigger guanosine tetraphosphate accumulation in Caulobacter crescentus. J Bacteriol 145:1463–1465.

33. Geiger T, Kästle B, Gratani FL, Goerke C, Wolz C. 2014. Two small (p)ppGpp synthases in Staphylococcus aureus mediate tolerance against cell envelope stress conditions. J Bacteriol 196:894–902.

34. Corrigan RM, Bellows LE, Wood A, Gründling A. 2016. ppGpp negatively impacts ribosome assembly affecting growth and antimicrobial tolerance in Gram-positive bacteria. Proc Natl Acad Sci U S A 113:E1710-9.

35. Hecker M, Pane-Farre J, Volker U. 2007. SigB-dependent general stress response in Bacillus subtilis and related gram-positive bacteria. Annu Rev Microbiol2007/11/24. 61:215–236.

36. Benson AK, Haldenwang WG. 1993. Regulation of sigma B levels and activity in Bacillus subtilis. J Bacteriol 175:2347–2356.

37. Völker U, Engelmann S, Maul B, Riethdorf S, Völker A, Schmid R, Mach H, Hecker M. 1994. Analysis of the induction of general stress proteins of Bacillus subtilis. Microbiology 140 (Pt 4:741–752.

38. Yang X, Kang CM, Brody MS, Price CW. 1996. Opposing pairs of serine protein kinases and phosphatases transmit signals of environmental stress to activate a bacterial transcription factor. Genes Dev 10:2265–2275.

39. Britos L, Abeliuk E, Taverner T, Lipton M, McAdams H, Shapiro L. 2011. Regulatory response to carbon starvation in Caulobacter crescentus. PLoS One2011/04/16. 6:e18179.

40. Fiebig A, Herrou J, Willett J, Crosson S. 2015. General Stress Signaling in the Alphaproteobacteria. Annu Rev Genet 49:603–625.

41. Staron A, Mascher T. 2010. General stress response in alpha-proteobacteria: PhyR and beyond. Mol Microbiol 78:271–277.

42. Francez-Charlot A, Kaczmarczyk A, Fischer HM, Vorholt JA. 2015. The general stress response in Alphaproteobacteria. Trends Microbiol 23:164–171.

43. Poindexter JS. 1964. Biological properties and classification of teh Caulobacter group. Bacteriol Rev 28:231–295.

44. Orcutt BN, Sylvan JB, Knab NJ, Edwards KJ. 2011. Microbial ecology of the dark ocean above, at, and below the seafloor. Microbiol Mol Biol Rev 75:361–422.

45. Portillo MC, Gonzalez JM. 2008. Microbial communities and immigration in volcanic environments of Canary Islands (Spain). Naturwissenschaften 95:307–315.

46. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J, Richardson TH, Noordewier M, Rappé MS, Short JM, Carrington JC, Mathur EJ. 2005. Genome Streamlining in a Cosmopolitan Oceanic Bacterium. Science (80-) 309:1242 LP – 1245.

47. Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe A, Idesawa K, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Mochizuki Y, Nakayama S, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C,

Yamada M, Tabata S. 2000. Complete Genome Structure of the Nitrogen-fixing Symbiotic Bacterium Mesorhizobium loti. DNA Res 7:331–338.

48. Li H, Tong Y, Huang Y, Bai J, Yang H, Liu W, Cao W. 2012. Complete genome sequence of Bartonella quintana, a bacterium isolated from rhesus macaques. J Bacteriol 194:6347.

49. Ettema TJG, Andersson SGE. 2009. The alpha-proteobacteria: the Darwin finches of the bacterial world. Biol Lett 5:429–432.

50. Philippot L, Andersson SGEE, Battin TJ, Prosser JI, Schimel JP, Whitman WB, Hallin S. 2010. The ecological coherence of high bacterial taxonomic ranks. Nat Rev Microbiol 8:523–529.

51. Batut J, Andersson SG, O'Callaghan D. 2004. The evolution of chronic infection strategies in the alpha-proteobacteria. Nat Rev Microbiol2004/11/20. 2:933–945.

52. Bruce D. 1887. Note on the discovery of a microorganism in Malta Fever. [John Brigg], [London].

53. Bang B. 1897. Die Aetiologie des seuchenhaften ("infectiösen") Verwerfens. Zietschrift für tiermedizin 1:241–278.

54. Seleem MN, Boyle SM, Sriranganathan N. 2010. Brucellosis: A re-emerging zoonosis. Vet Microbiol 140:392–398.

55. Boschiroli ML, Foulongne V, O'Callaghan D. 2001. Brucellosis: A worldwide zoonosis. Curr Opin Microbiol 4:58–64.

56. Galińska EM, Zagórski J. 2013. Brucellosis in humans - Etiology, diagnostics, clinical forms. Ann Agric Environ Med 20:233–238.

57. Ridler AL, West DM. 2011. Control of Brucella ovis Infection in Sheep. Vet Clin North Am - Food Anim Pract 27:61–66.

58. Godfroid J, Garin-Bastuji B, Saegerman C, Blasco JM. 2013. Brucellosis in terrestrial wildlife. OIE Rev Sci Tech 32:27–42.

59. Moreno E. 2014. Retrospective and prospective perspectives on zoonotic brucellosis. Front Microbiol 5:1–18.

60. Whatmore AM. 2009. Current understanding of the genetic diversity of Brucella, an expanding genus of zoonotic pathogens. Infect Genet Evol 9:1168–1184.

61. Alton GG, Forsy JRL. 1996. Brucella, p. . *In* Baron, S (ed.), Medical Microbiology. Galveston (TX).

62. Franco MP, Mulder M, Gilman RH, Smits HL. 2007. Human brucellosis. Lancet Infect Dis 7:775–786.

63. Al Dahouk S, Hofer E, Tomaso H, Vergnaud G, Le Flèche P, Cloeckaert A, Koylass MS, Whatmore AM, Nöckler K, Scholz HC. 2012. Intraspecies diodiversity of the genetically homologous species Brucella microti. Appl Environ Microbiol 78:1534–1543.

64. Spink WW, Morisset R. 1970. Epidemic canine brucellosis due to a new species: Brucella canis. Trans Am Clin Climatol Assoc 81:43–50.

65. Buddle M, Boyes B. 1953. A Brucella mutant causing genital disease of sheep in New

Zealand. Aust Vet J 29:145–153.

66. STOENNER HG, LACKMAN DB. 1957. A new species of Brucella isolated from the desert wood rat, Neotoma lepida Thomas. Am J Vet Res 18:947–951.

67. Foster G, Jahans KL, Reid RJ, Ross HM. 1996. Isolation of Brucella species from cetaceans, seals and an otter. Vet Rec 138:583–586.

68. Ross HM, Foster G, Reid RJ, Jahans KL, MacMillan AP. 1994. Brucella species infection in sea-mammals. Vet Rec. England.

69. Ewalt DR, Payeur JB, Martin BM, Cummins DR, Miller WG. 1994. Characteristics of a Brucella species from a bottlenose dolphin (Tursiops truncatus). J Vet diagnostic Investig Off Publ Am Assoc Vet Lab Diagnosticians, Inc 6:448–452.

70. Hubálek Z, Scholz HC, Sedlácek I, Melzer F, Sanogo YO, Nesvadbová J. 2007. Brucellosis of the common vole (Microtus arvalis). Vector Borne Zoonotic Dis 7:679–687.

71. Scholz HC, Nöckler K, Göllner C, Bahn P, Vergnaud G, Tomaso H, Al Dahouk S, Kämpfer P, Cloeckaert A, Maquart M, Zygmunt MS, Whatmore AM, Pfeffer M, Huber B, Busse H-J, De BK. 2010. Brucella inopinata sp. nov., isolated from a breast implant infection. Int J Syst Evol Microbiol 60:801–808.

72. Whatmore AM, Davison N, Cloeckaert A, Al Dahouk S, Zygmunt MS, Brew SD, Perrett LL, Koylass MS, Vergnaud G, Quance C, Scholz HC, Dick EJ, Hubbard G, Schlabritz-Loutsevitch NE. 2014. Brucella papionis sp. nov., isolated from baboons (Papio spp.). Int J Syst Evol Microbiol 64:4120–4128.

73. Scholz HC, Revilla-Fernández S, Dahouk S Al, Hammerl JA, Zygmunt MS, Cloeckaert A, Koylass M, Whatmore AM, Blom J, Vergnaud G, Witte A, Aistleitner K, Hofer E. 2016. Brucella vulpis sp. nov., isolated from mandibular lymph nodes of red foxes (Vulpes vulpes). Int J Syst Evol Microbiol 66:2090–2098.

74. Soler-Lloréns PF, Quance CR, Lawhon SD, Stuber TP, Edwards JF, Ficht TA, Robbe-Austerman S, O'Callaghan D, Keriel A. 2016. A Brucella spp. isolate from a Pac-Man frog (Ceratophrys ornata) reveals characteristics departing from classical brucellae. Front Cell Infect Microbiol 6:1–16.

75. Al Dahouk S, Köhler S, Occhialini A, Jiménez de Bagüés MP, Hammerl JA, Eisenberg T, Vergnaud G, Cloeckaert A, Zygmunt MS, Whatmore AM, Melzer F, Drees KP, Foster JT, Wattam AR, Scholz HC. 2017. Brucella spp. of amphibians comprise genomically diverse motile strains competent for replication in macrophages and survival in mammalian hosts. Sci Rep 7:44420.

76. Fiebig A, Vrentas CE, Le T, Huebner M, Boggiatto PM, Olsen SC, Crosson S. 2021. Quantification of Brucella abortus population structure in a natural host. Proc Natl Acad Sci U S A 118.

77. 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 65:523–541.

78. de Figueiredo P, Ficht TA, Rice-Ficht A, Rossetti CA, Adams LG. 2015. Pathogenesis and immunobiology of brucellosis: review of Brucella-host interactions. Am J Pathol 185:1505–1517.

79. Tuon FF, Gondolfo RB, Cerchiari N. 2017. Human-to-human transmission of Brucella – a systematic review. Trop Med Int Heal 22:539–546.

80. Tsolis RM, Seshadri R, Santos RL, Sangari FJ, García 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 One2018/05/23. 4:https://doi:10.1371/journal.pone.0005519.

81. De Jong MF, Rolán HG, Tsolis RM. 2010. Microreview: Innate immune encounters of the (Type) 4th kind: Brucella. Cell Microbiol 12:1195–1202.

82. Tsolis RM, Young GM, Solnick J V., Bäumler AJ. 2008. From bench to bedside: Stealth of enteroinvasive pathogens. Nat Rev Microbiol 6:883–892.

83. Ahmed W, Zheng K, Liu ZF. 2016. Establishment of chronic infection: Brucella's stealth strategy. Front Cell Infect Microbiol 6:1–12.

84. Teng TS, Ji AL, Ji XY, Li YZ. 2017. Neutrophils and immunity: From bactericidal action to being conquered. J Immunol Res 2017.

85. Mora-Cartín R, Gutiérrez-Jiménez C, Alfaro-Alarcón A, Chaves-Olarte E, Chacón-Díaz C, Barquero-Calvo E, Moreno E. 2019. Neutrophils Dampen Adaptive Immunity in Brucellosis. Infect Immun 87.

86. Ben-Tekaya H, Gorvel J-P, Dehio C. 2013. Bartonella and Brucella--weapons and strategies for stealth attack. Cold Spring Harb Perspect Med 3.

87. Lu YC, Yeh WC, Ohashi PS. 2008. LPS/TLR4 signal transduction pathway. Cytokine 42:145–151.

88. Conde-Alvarez R, Arce-Gorvel V, Iriarte M, Mancek-Keber M, Barquero-Calvo E, Palacios-Chaves L, Chacon-Diaz C, Chaves-Olarte E, Martirosyan A, von Bargen K, Grillo MJ, Jerala R, Brandenburg K, Llobet E, Bengoechea JA, Moreno E, Moriyon I, Gorvel JP. 2012. The lipopolysaccharide core of Brucella abortus acts as a shield against innate immunity recognition. PLoS Pathog 8:e1002675.

89. Weiss DS, Takeda K, Akira S, Zychlinsky A, Moreno E. 2005. MyD88, but not toll-like receptors 4 and 2, is required for efficient clearance of Brucella abortus. Infect Immun 73:5137–5143.

90. Kawasaki T, Kawai T. 2014. Toll-Like Receptor Signaling Pathways . Front Immunol .

91. Andersen-Nissen E, Smith KD, Strobe KL, Barrett SLR, Cookson BT, Logan SM, Aderem A. 2005. Evasion of Toll-like receptor 5 by flagellated bacteria. Proc Natl Acad Sci U S A 102:9247–9252.

92. Letesson JJ, Lestrate P, Delrue RM, Danese I, Bellefontaine F, Fretin D, Taminiau B, Tibor A, Dricot A, Deschamps C, Haine V, Leonard S, Laurent T, Mertens P, Vandenhaute J, De Bolle X. 2002. Fun stories about Brucella: The "furtive nasty bug." Vet Microbiol 90:317–328.

93. Fretin D, Fauconnier A, Köhler S, Halling S, Léonard S, Nijskens C, Ferooz J, Lestrate P, Delrue RM, Danese I, Vandenhaute J, Tibor A, DeBolle X, Letesson JJ. 2005. The sheathed

flagellum of Brucella melitensis is involved in persistence in a murine model of infection. Cell Microbiol 7:687–698.

94. Covert J, Mathison AJ, Eskra L, Banai M, Splitter G. 2009. Brucella melitensis, B. neotomae and B. ovis elicit common and distinctive macrophage defense transcriptional responses. Exp Biol Med 234:1450–1467.

95. Celli J, De Chastellier C, Franchini DM, Pizarro-Cerda J, Moreno E, Gorvel JP. 2003. Brucella evades macrophage killing via VirB-dependent sustained interactions with the endoplasmic reticulum. J Exp Med 198:545–556.

96. Celli J. 2015. The changing nature of the Brucella-containing vacuole. Cell Microbiol 17:951–958.

97. Celli J. 2019. The Intracellular Life Cycle of Brucella spp. Bact Intracellularity 7:101–111.

98. Sá JC, Silva TMA, Costa ÉA, Silva APC, Tsolis RM, Paixão TA, Carvalho Neta A V., Santos RL. 2012. The virB-encoded type IV secretion system is critical for establishment of infection and persistence of Brucella ovis infection in mice. Vet Microbiol 159:130–140.

99. Delrue RM, Martinez-Lorenzo M, Lestrate P, Danese I, Bielarz V, Mertens P, De Bolle X, Tibor A, Gorvel JP, Letesson JJ. 2001. Identification of Brucella spp. genes involved in intracellular trafficking. Cell Microbiol 3:487–497.

100. Starr T, Child R, Wehrly TD, Hansen B, Hwang S, López-Otin C, Virgin HW, Celli J. 2012. Selective subversion of autophagy complexes facilitates completion of the Brucella intracellular cycle. Cell Host Microbe 11:33–45.

101. Roop RM, Gaines JM, Anderson ES, Caswell CC, Martin DW. 2009. Survival of the fittest: How Brucella strains adapt to their intracellular niche in the host. Med Microbiol Immunol 198:221–238.

102. Roop II RM, Gee JM, Robertson GT, Richardson JM, Ng W-L, Winkler ME. 2003. Brucella stationary-phase gene expression and virulence. Annu Rev Microbiol 57:57–76.

103. Jiang X, Leonard B, Benson R, Baldwin CL. 1993. Macrophage control of Brucella abortus: Role of reactive oxygen intermediates and nitric oxide. Cell Immunol 151:309–319.

104. Muraille E, Leo O, Moser M. 2014. Th1/Th2 paradigm extended: Macrophage polarization as an unappreciated pathogen-driven escape mechanism? Front Immunol 5:1–12.

105. Brundu S FA. 2015. Polarization and repolarization of macrophages. J Clin Cell Immunol 06:1–10.

106. Reniere ML. 2018. Reduce, Induce, Thrive: Bacterial Redox Sensing during Pathogenesis. J Bacteriol 200:1–27.

107. Robertson GT, Roop RM. 1999. The Brucella abortus host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. Mol Microbiol 34:690–700.

108. Hanna N, Ouahrani-Bettache S, Drake KL, Adams LG, Köhler S, Occhialini A. 2013. Global Rsh-dependent transcription profile of Brucella suis during stringent response unravels adaptation to nutrient starvation and cross-talk with other stress responses. BMC Genomics 14:459.

109. Yang P, Huang S, Yan X, Huang G, Dong X, Zheng T, Yuan D, Wang R, Li R, Tan Y, Xu A. 2014. Origin of the phagocytic respiratory burst and its role in gut epithelial phagocytosis in a basal chordate. Free Radic Biol Med 70:54–67.

110. Halliwell B, Gutteridge JM. 1984. Oxygen toxicity, oxygen radicals, transition metals and disease. Biochem J 219:1–14.

111. Ezraty B, Gennaris A, Barras F, Collet JF. 2017. Oxidative stress, protein damage and repair in bacteria. Nat Rev Microbiol 15:385–396.

112. Phaniendra A, Jestadi DB, Periyasamy L. 2015. Free radicals: properties, sources, targets, and their implication in various diseases. Indian J Clin Biochem 30:11–26.

113. Imlay JA. 2013. The molecular mechanisms and physiological consequences of oxidative stress: Lessons from a model bacterium. Nat Rev Microbiol 11:443–454.

114. McCord JM, Fridovich I. 1969. Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). J Biol Chem 244:6049–6055.

115. Aust SD, Morehouse LA, Thomas CE. 1985. Role of metals in oxygen radical reactions. J Free Radic Biol Med 1:3–25.

116. Chelikani P, Fita I, Loewen PC. 2004. Diversity of structures and properties among catalases. Cell Mol Life Sci 61:192–208.

117. Brigelius-Flohé R, Maiorino M. 2013. Glutathione peroxidases. Biochim Biophys Acta 1830:3289–3303.

118. Rhee SG, Kang SW, Chang TS, Jeong W, Kim K. 2001. Peroxiredoxin, a novel family of peroxidases. IUBMB Life 52:35–41.

119. Forrester MT, Foster MW. 2012. Protection from nitrosative stress: A central role for microbial flavohemoglobin. Free Radic Biol Med 52:1620–1633.

120. Gardner AM, Helmick RA, Gardner PR. 2002. Flavorubredoxin, an inducible catalyst for nitric oxide reduction and detoxification in Escherichia coli. J Biol Chem 277:8172–8177.

121. Winterbourn CC, Hampton MB. 2008. Thiol chemistry and specificity in redox signaling. Free Radic Biol Med 45:549–561.

122. Vogt W. 1995. Oxidation of methionyl residues in proteins: tools, targets, and reversal. Free Radic Biol Med 18:93–105.

123. Meister A, Anderson ME. 1983. Glutathione. Annu Rev Biochem 52:711–760.

124. Masip Ll, Veeravalli K, Georgiou G. 2006. The many faces of glutathione in bacteria. Antioxidants Redox Signal 8:753–762.

125. Kerksick C, Willoughby D. 2005. The antioxidant role of glutathione and N-acetyl-cysteine supplements and exercise-induced oxidative stress. J Int Soc Sports Nutr 2:38–44.

126. Park S, Imlay JA. 2003. High levels of intracellular cysteine promote oxidative DNA damage by driving the Fenton reaction. J Bacteriol 185:1942–1950.

127. Reddy VN. 1990. Glutathione and its function in the lens--an overview. Exp Eye Res 50:771–778.

128. Fahey RC. 2013. Glutathione analogs in prokaryotes. Biochim Biophys Acta 1830:3182–3198.

129. Winterbourn CC, Kettle AJ. 2013. Redox reactions and microbial killing in the neutrophil phagosome. Antioxid Redox Signal 18:642–660.

130. Spitznagel JK. 1977. Bactericidal mechanisms of the granulocyte. Prog Clin Biol Res 13:103–131.

131. Fang FC. 2004. Antimicrobial reactive oxygen and nitrogen species: Concepts and controversies. Nat Rev Microbiol 2:820–832.

132. Storz G, Tartaglia LA. 1992. OxyR: a regulator of antioxidant genes. J Nutr 122:627–630.

133. Dubbs JM, Mongkolsuk S. 2012. Peroxide-Sensing Transcriptional Regulators in Bacteria. J Bacteriol 194:5495 LP – 5503.

134. Duarte V, Latour J-M. 2010. PerR vs OhrR: selective peroxide sensing in Bacillus subtilis. Mol Biosyst 6:316–323.

135. Fritsch VN, Loi V Van, Busche T, Tung QN, Lill R, Horvatek P, Wolz C, Kalinowski J, Antelmann H. 2020. The alarmone (p)ppGpp confers tolerance to oxidative stress during the stationary phase by maintenance of redox and iron homeostasis in Staphylococcus aureus. Free Radic Biol Med 161:351–364.

136. Gee JM, Valderas MW, Kovach ME, Grippe VK, Robertson GT, Ng W-L, Richardson JM, Winkler ME, Roop 2nd RM. 2005. The Brucella abortus Cu,Zn superoxide dismutase is required for optimal resistance to oxidative killing by murine macrophages and wild-type virulence in experimentally infected mice. Infect Immun 73:2873–2880.

137. Martin DW, Baumgartner JE, Gee JM, Anderson ES, Martin Roop I. 2012. SodA is a major metabolic antioxidant in Brucella abortus 2308 that plays a significant, but limited, role in the virulence of this strain in the mouse model. Microbiol (United Kingdom) 158:1767–1774.

138. Lensmire JM, Hammer ND. 2019. Nutrient sulfur acquisition strategies employed by bacterial pathogens. Curr Opin Microbiol 47:52–58.

139. Spyrakis F, Singh R, Cozzini P, Campanini B, Salsi E, Felici P, Raboni S, Benedetti P, Cruciani G, Kellogg GE, Cook PF, Mozzarelli A. 2013. Isozyme-specific ligands for O-acetylserine sulfhydrylase, a novel antibiotic target. PLoS One 8:e77558.

140. Schnell R, Sriram D, Schneider G. 2015. Pyridoxal-phosphate dependent mycobacterial cysteine synthases: Structure, mechanism and potential as drug targets. Biochim Biophys Acta - Proteins Proteomics 1854:1175–1183.

141. Gebhardt MJ, Gallagher LA, Jacobson RK, Usacheva EA, Peterson LR, Zurawski D V, Shuman HA. 2015. Joint Transcriptional Control of Virulence and Resistance to Antibiotic and Environmental Stress in Acinetobacter baumannii. MBio 6:e01660-15.

142. Gebhardt MJ, Czyz DM, Singh S, Zurawski D V, Becker L, Shuman HA. 2020. GigC, a LysR Family Transcription Regulator, Is Required for Cysteine Metabolism and Virulence in Acinetobacter baumannii Infect Immun 89:e00180-20.

143. Lithgow JK, Hayhurst EJ, Cohen G, Aharonowitz Y, Foster SJ. 2004. Role of a Cysteine Synthase in Staphylococcus aur. J Bacteriol 186:1579 LP – 1590.

144. Lensmire JM, Dodson JP, Hsueh BY, Wischer MR, Delekta PC, Shook JC, Ottosen EN, Kies PJ, Ravi J, Hammer ND. 2020. The Staphylococcus aureus Cystine Transporters TcyABC and TcyP Facilitate Nutrient Sulfur Acquisition during Infection. Infect Immun 88.

145. Connolly JP, Comerci D, Alefantis TG, Walz A, Quan M, Chafin R, Grewal P, Mujer C V, Ugalde RA, DelVecchio VG. 2006. Proteomic analysis of Brucella abortus cell envelope and identification of immunogenic candidate proteins for vaccine development. Proteomics 6:3767–3780.

146. Jain S, Afley P, Kumar S. 2013. Immunological responses to recombinant cysteine synthase A of Brucella abortus in BALB/c mice. World J Microbiol Biotechnol 29:907–913.

147. On S, Ovis B, Disease G, Sheep OF, New IN, Buddle BYMB. 1955. Genital Disease of Sheep in New Zealand.

148. Poltorak A, Smirnova I, He X, Liu MY, Van Huffel C, McNally O, Birdwell D, Alejos E, Silva M, Du X, Thompson P, Chan EK, Ledesma J, Roe B, Clifton S, Vogel SN, Beutler B. 1998. Genetic and physical mapping of the Lps locus: identification of the toll-4 receptor as a candidate gene in the critical region. Blood Cells Mol Dis 24:340–355.

149. Murdock JL, Núñez G. 2016. TLR4: The Winding Road to the Discovery of the LPS Receptor. J Immunol 197:2561–2562.

150. Liu Y, Sun J, Peng X, Dong H, Qin Y, Shen Q, Jiang H, Xu G, Feng Y, Sun S, Ding J, Chen R. 2020. Deletion of the LuxR-type regulator VjbR in Brucella canis affects expression of type IV secretion system and bacterial virulence, and the mutant strain confers protection against Brucella canis challenge in mice. Microb Pathog 139:103865.

151. Li P, Tian M, Bao Y, Hu H, Liu J, Yin Y, Ding C, Wang S, Yu S. 2017. Brucella rough mutant induce macrophage death via activating IRE1a pathway of endoplasmic reticulum stress by enhanced T4SS secretion. Front Cell Infect Microbiol 7:1–15.

152. Pei J, Ficht TA. 2004. Brucella abortus rough mutants are cytopathic for macrophages in culture. Infect Immun 72:440–450.

153. Herrou J, Willett JW, Fiebig A, Czyż 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. bioRxiv 201:1–20.

154. CFSPH. 2007. Ovine Epididymitis : Brucella ovis. CFSPH Publ 1–4.

155. Antunes JMA de P, Allendorf SD, Appolinário CM, Cagnini DQ, Figueiredo PR, Júnior JB, Baños JV, Kocan KM, de la Fuente J, Megid J. 2013. Rough virulent strain of Brucella ovis induces pro- and anti-inflammatory cytokines in reproductive tissues in experimentally infected rams. Vet Microbiol 161:339–343.

156. Gouletsou PG, Fthenakis GC. 2015. Microbial diseases of the genital system of rams or bucks. Vet Microbiol 181:130–135.

157. 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 11:1–7.

158. Alton GG, Jones LM, Pietz DE. 1975. Laboratory techniques in brucellosis. Monogr Ser World Heal Organ 1–163.

159. Meyer ME. 1969. Phenotypic comparison of Brucella ovis to the DNA-homologous Brucella species. Am J Vet Res 30:1757–1764.

160. Petersen E, Rajashekara G, Sanakkayala N, Eskra L, Harms J, Splitter G. 2013. Erythritol triggers expression of virulence traits in Brucella melitensis. Microbes Infect2013/02/16. 15:440–449.

161. Varesio LM, Willett JW, Fiebig A, Crosson S. 2019. A carbonic anhydrase pseudogene sensitizes select Brucella lineages to low CO2 tension. J Bacteriol https://doi.org/10.1128/JB.00509-19.

162. Pérez-Etayo L, De Miguel MJ, Conde-Álvarez R, Muñoz PM, Khames M, Iriarte M, Moriyón I, Zúñiga-Ripa A. 2018. The CO 2 -dependence of Brucella ovis and Brucella abortus biovars is caused by defective carbonic anhydrases. Vet Res 49:1–12.

163. Rambow-Larsen AA, Petersen EM, Gourley CR, Splitter GA. 2009. Brucella regulators: self-control in a hostile environment. Trends Microbiol 17:371–377.

164. Henry B, Traum J, Haring C. Methods for the isolation of Brucella abortus. Hilgardia 6:355–379.

165. Huddleson I. 1921. The importance of an increased carbon dioxide tension in growing Bact. abortus (Bang). Cornell Vet 11:201–215.

166. Smith T. 1924. SOME CULTURAL CHARACTERS OF BACILLUS ABORTUS (BANG) WITH SPECIAL REFERENCE TO CO(2) REQUIREMENTS. J Exp Med 40:219–232.

167. Wilson GS. 1931. The Gaseous Requirements of Br. abortus (Bovine type). Br J Exp Pathol 12:88–92.

168. NEWTON JW, MARR AG, WILSON JB. 1954. Fixation of C14O2 into nucleic acid constituents by Brucella abortus. J Bacteriol 67:233–236.

169. MARR AG, WILSON JB. 1951. Fixation of C14O2 in amino acids by Brucella abortus. Arch Biochem Biophys 34:442–448.

170. TEPPER BS, WILSON JB. 1958. Fixation and distribution of C14O2 in Brucella abortus. J Bacteriol 76:24–28.

171. GERHARDT P, WILSON JB. 1950. Attempts to replace the added carbon dioxide required by some strains of Brucella abortus. J Bacteriol 59:311–312.

172. Spink W. 1956. The Evolution of the Concept that Brucellosis is a Disease of Animals and Man, p. 3–28. *In* University of Minnesota (ed.), The Nature of Brucellosis. Minnapolis, MN.

173. Evans A. 1918. Further studies on bacterium abortus and related bacteria. J Infect Dis 22:580–593.

174. Evans AC. 1923. The Nomenclature of the Melitensis-Abortus Group of Bacterial Organisms. Public Heal Reports 38:1943–1948.

175. Hardy A, CF J, Borts I, Hardy G. 1931. Undulant Fever with Special Reference to a Study of Brucella Infection in Iowa. J Am Med Assoc 97:52.

176. Fitch CP. 1922. The Cultivation of Bacterium Abortus Bang. J Infect Dis 31:233–236.

177. BUDDLE MB. 1956. Studies on Brucella ovis (n.sp.), a cause of genital disease of sheep in New Zealand and Australia. J Hyg (Lond) 54:351–364.

178. Marr AG, Wilson JB. 1950. The Carbon Dioxide Requirements of Brucella Abortus, p. 151–155. *In* Thrid Inter-American Congress on Brucellosis. The National Academy Press.

179. Blombach B, Takors R. 2015. CO2 - intrinsic product, essential substrate, and regulatory trigger of microbial and mammalian production processes. Front Bioeng Biotechnol 3:1–11.

180. Erb TJ. 2011. Carboxylases in natural and synthetic microbial pathways. Appl Environ Microbiol 77:8466–8477.

181. Fuchs G. 2011. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? Annu Rev Microbiol 65:631–658.

182. Tong L. 2013. Structure and function of biotin-dependent carboxylases. Cell Mol Life Sci 70:863–891.

183. Waldrop GL, Holden HM, St Maurice M. 2012. The enzymes of biotin dependent CO_2 metabolism: what structures reveal about their reaction mechanisms. Protein Sci 21:1597–1619.

184. Meldrum NU, Roughton FJ. 1933. Carbonic anhydrase. Its preparation and properties. J Physiol 80:113–142.

185. Ogawa T, Noguchi K, Saito M, Nagahata Y, Kato H, Ohtaki A, Nakayama H, Dohmae N, Matsushita Y, Odaka M, Yohda M, Nyunoya H, Katayama Y. 2013. Carbonyl Sulfide Hydrolase from Thiobacillus thioparus Strain THI115 Is One of the β-Carbonic Anhydrase Family Enzymes. J Am Chem Soc 135:3818–3825.

186. Smeulders MJ, Barends TRM, Pol A, Scherer A, Zandvoort MH, Udvarhelyi A, Khadem AF, Menzel A, Hermans J, Shoeman RL, Wessels HJCT, van den Heuvel LP, Russ L, Schlichting I, Jetten MSM, Op den Camp HJM. 2011. Evolution of a new enzyme for carbon disulphide conversion by an acidothermophilic archaeon. Nature 478:412–416.

187. Smith KS, Ferry JG. 2000. Prokaryotic carbonic anhydrases. FEMS Microbiol Rev 24:335–366.

188. Supuran CT, Capasso C. 2017. An Overview of the Bacterial Carbonic Anhydrases. Metabolites 7.

189. Tripp BC, Smith K, Ferry JG. 2001. Carbonic anhydrase: new insights for an ancient enzyme. J Biol Chem 276:48615–48618.

190. Supuran CT. 2016. Structure and function of carbonic anhydrases. Biochem J 473:2023–2032.

191. Sly WS. 1995. Human Carbonic Anhydrases and Carbonic Anhydrase Deficiencies. Annu Rev Biochem 64:375–401.

192. Briganti F, Mangani S, Scozzafava A, Vernaglione G, Supuran CT. 1999. Carbonic anhydrase catalyzes cyanamide hydration to urea: is it mimicking the physiological reaction? JBIC J Biol Inorg Chem 4:528–536.

193. Henry RP. 1996. Multiple Roles of Carbonic Anhydrase in Cellular Transport and Metabolism. Annu Rev Physiol 58:523–538.

194. 1973. Carbonic Anhydrase and Bone Resorption. Nutr Rev 31:101–102.

195. Dobyan DC, Bulger RE. 1982. Renal carbonic anhydrase. Am J Physiol Physiol 243:F311–F324.

196. Aspatwar A, Haapanen S, Parkkila S. 2018. An Update on the Metabolic Roles of Carbonic Anhydrases in the Model Alga Chlamydomonas reinhardtii. Metabolites 8:22.

197. Moroney J V, Ma Y, Frey WD, Fusilier KA, Pham TT, Simms TA, DiMario RJ, Yang J, Mukherjee B. 2011. The carbonic anhydrase isoforms of Chlamydomonas reinhardtii: intracellular location, expression, and physiological roles. Photosynth Res 109:133–149.

198. Badger MR, Price GD. 1994. The Role of Carbonic Anhydrase in Photosynthesis. Annu Rev Plant Physiol Plant Mol Biol 45:369–392.

199. Matsuda Y, Hopkinson BM, Nakajima K, Dupont CL, Tsuji Y. 2017. Mechanisms of carbon dioxide acquisition and CO(2) sensing in marine diatoms: a gateway to carbon metabolism. Philos Trans R Soc London Ser B, Biol Sci 372.

200. Jensen EL, Maberly SC, Gontero B. 2020. Insights on the Functions and Ecophysiological Relevance of the Diverse Carbonic Anhydrases in Microalgae. Int J Mol Sci .

201. Krungkrai J, Krungkrai SR, Supuran CT. 2008. Carbonic anhydrase inhibitors: inhibition of Plasmodium falciparum carbonic anhydrase with aromatic/heterocyclic sulfonamides-in vitro and in vivo studies. Bioorg Med Chem Lett 18:5466–5471.

202. Reyes P, Rathod PK, Sanchez DJ, Mrema JEK, Rieckmann KH, Heidrich H-G. 1982. Enzymes of purine and pyrimidine metabolism from the human malaria parasite, Plasmodium falciparum. Mol Biochem Parasitol 5:275–290.

203. Reungprapavut S, Krungkrai SR, Krungkrai J. 2004. Plasmodium falciparum Carbonic Anhydrase is a Possible Target for Malaria Chemotherapy. J Enzyme Inhib Med Chem 19:249–256.

204. Fan S-H, Ebner P, Reichert S, Hertlein T, Zabel S, Lankapalli AK, Nieselt K, Ohlsen K, Götz F. 2019. MpsAB is important for Staphylococcus aureus virulence and growth at atmospheric CO(2) levels. Nat Commun 10:3627.

205. Valdivia RH, Falkow S. 1997. Fluorescence-Based Isolation of Bacterial Genes Expressed Within Host Cells. Science (80-) 277:2007 LP – 2011.

206. Merlin C, Masters M, McAteer S, Coulson A. 2003. Why is carbonic anhydrase essential to Escherichia coli? J Bacteriol 185:6415–6424.

207. Joseph P, Turtaut F, Ouahrani-Bettache S, Montera JL, Nishimori I, Minakuchi T, Vullo D, Scozzafava A, Köhler S, Winum JY, Supuran CT. 2010. Cloning, characterization, and

inhibition studies of a β-carbonic anhydrase from brucella suis. J Med Chem 53:2277–2285.

208. Köhler 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.

209. Herrou J, Crosson S. 2013. Molecular structure of the Brucella abortus metalloprotein RicA, a Rab2-binding virulence effector. Biochemistry 52:9020–9028.

210. Ombouma J, Vullo D, Köhler S, Dumy P, Supuran CT, Winum JY. 2015. N-glycosyl-N-hydroxysulfamides as potent inhibitors of Brucella suis carbonic anhydrases. J Enzyme Inhib Med Chem 30:1010–1012.

211. Demerec M, Adelberg EA, Clark AJ, Hartman PE. 1966. A proposal for a uniform nomenclature in bacterial genetics. Genetics 54:61–76.

212. DiMario RJ, Clayton H, Mukherjee A, Ludwig M, Moroney J V. 2017. Plant Carbonic Anhydrases: Structures, Locations, Evolution, and Physiological Roles. Mol Plant2016/09/16. 10:30–46.

213. Supuran CT. 2016. Legionella pneumophila carbonic anhydrases: Underexplored antibacterial drug targets. Pathogens 5:24–29.

214. Joseph P, Ouahrani-Bettache S, Montero JL, Nishimori I, Minakuchi T, Vullo D, Scozzafava A, Winum JY, Köhler S, Supuran CT. 2011. A new β-carbonic anhydrase from Brucella suis, its cloning, characterization, and inhibition with sulfonamides and sulfamates, leading to impaired pathogen growth. Bioorganic Med Chem 19:1172–1178.

215. Alton GG, Jones LM, Pietz DE, Organization WH, Nations F and AO of the U. 1975. Laboratory techniques in brucellosis / G. G. Alton, Lois M. Jones & D. E. Pietz2nd ed. World Health Organization, Geneva PP - Geneva.

216. Paul BJ, Ross W, Gaal T, Gourse RL. 2004. rRNA transcription in Escherichia coli. Annu Rev Genet 38:749–770.

217. Rao NN, Gómez-García MR, Kornberg A. 2009. Inorganic polyphosphate: essential for growth and survival. Annu Rev Biochem 78:605–647.

218. Gray MJ, Jakob U. 2015. Oxidative stress protection by polyphosphate--new roles for an old player. Curr Opin Microbiol 24:1–6.

219. Henry JT, Crosson S. 2013. Chromosome replication and segregation govern the biogenesis and inheritance of inorganic polyphosphate granules. Mol Biol Cell 24:3177–3186.

220. 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 200:1–17.

221. 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 188:7707–7710.

222. Kim HS, Caswell CC, Foreman R, Roop RM, Crosson S. 2013. The Brucella abortus general stress response system regulates chronic mammalian infection and is controlled by phosphorylation and proteolysis. J Biol Chem 288:13906–13916.

223. 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.

224. Li WH, Gojobori T, Nei M. 1981. Pseudogenes as a paradigm of neutral evolution. Nature 292:237–239.

225. Petrov DA, Hartl DL. 2000. Pseudogene evolution and natural selection for a compact genome. J Hered 91:221–227.

226. Kuo C-H, Ochman H. 2010. The extinction dynamics of bacterial pseudogenes. PLoS Genet 6.

227. Corbel MJ, Hendry DM. 1985. Urease activity of Brucella species. Res Vet Sci 38:252–253.

228. Sangari FJ, Seoane A, Rodríguez MC, Agüero J, Lobo JMG. 2007. Characterization of the urease operon of Brucella abortus and assessment of its role in virulence of the bacterium. Infect Immun 75:774–780.

229. Hong PC, Tsolis RM, Ficht TA. 2000. Identification of genes required for chronic persistence of Brucella abortus in mice. Infect Immun 68:4102–4107.

230. Macedo AA, Silva APC, Mol JPS, Costa LF, Garcia LNN, Araújo MS, Filho OAM, Paixão 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 10:1–16.

231. Sieira R, Comerci DJ, Sánchez 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 182:4849–4855.

232. Anand A, Olson CA, Yang L, Sastry A V., Catoiu E, Choudhary KS, Phaneuf P V., 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.

233. 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:e00365-18.

234. Gengenbacher M, Kaufmann SHE. 2012. Mycobacterium tuberculosis: success through dormancy. FEMS Microbiol Rev 36:514–532.

235. Free MJ, Schluntz GA, Jaffe RA. 1976. Respiratory gas tensions in tissues and fluids of the male rat reproductive tract. Biol Reprod 14:481–488.

236. Wenger RH, Katschinski DM. 2005. The hypoxic testis and post-meiotic expression of PAS domain proteins. Semin Cell Dev Biol 16:547–553.

237. Anderson TD, Cheville NF, Meador VP. 1986. Pathogenesis of Placentitis in the Goat Inoculated with Brucella abortus. II. Ultrastructural Studies. Vet Pathol 23:227–239.

238. 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.

239. Boutte CC, Crosson S. 2013. Bacterial lifestyle shapes stringent response activation. Trends Microbiol 21:174–180.

240. Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol Biol 1151:165–

188.

241. 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.

242. Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30.

243. 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 6:1–15.

244. Herrou J, Willett JW, Fiebig A, Varesio LM, Czyż DM, Cheng JX, Ultee E, Briegel A, Bigelow L, Babnigg G, Kim Y, Crosson S, Czyz DM, Cheng JX, Ultee E, Briegel A, Bigelow L, Babnigg G, Kim Y, Crosson S, Czyż DM, Cheng JX, Ultee E, Briegel A, Bigelow L, Babnigg G, Kim Y, Crosson S, 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.

245. Porte F, Liautard JP, Köhler S. 1999. Early acidification of phagosomes containing Brucella suis is essential for intracellular survival in murine macrophages. Infect Immun 67:4041–4047.

246. Kim HS, Willett JW, Jain-Gupta N, Fiebig A, Crosson S. 2014. The Brucella abortus virulence regulator, LovhK, is a sensor kinase in the general stress response signalling pathway. Mol Microbiol 94:913–925.

247. Rossetti CA, Galindo CL, Lawhon SD, Garner HR, Adams LG. 2009. Brucella melitensis global gene expression study provides novel information on growth phase-specific gene regulation with potential insights for understanding Brucella:host initial interactions. BMC Microbiol 9:1–14.

248. Kredich NM, Tomkins GM. 1966. The enzymic synthesis of L-cysteine in Escherichia coli and Salmonella typhimurium. J Biol Chem 241:4955–4965.

249. Sidhu-Muñoz RS, Sancho P, Vizcaíno N. 2018. Evaluation of human trophoblasts and ovine testis cell lines for the study of the intracellular pathogen Brucella ovis. FEMS Microbiol Lett 365:1–9.

250. Keer J, Smeulders MJ, Williams HD. 2001. A purF mutant of Mycobacterium smegmatis has impaired survival during oxygen-starved stationary phase. Microbiology 147:473–481.

251. Samant S, Lee H, Ghassemi M, Chen J, Cook JL, Mankin AS, Neyfakh AA. 2008. Nucleotide biosynthesis is critical for growth of bacteria in human blood. PLoS Pathog doi:10. 1371/journal.ppat.0040037.

252. Shaffer C, Guckes KR, Breland EJ, Floyd KA, Casella DP, Algood HMS, Clayton DB. 2017. Purine biosynthesis metabolically constrains intracellular survival of uropathogenic Escherichia coli. Infect Immun 85:1–14.

253. de Crécy-Lagard V, Jaroch M. 2020. Functions of bacterial tRNA modifications: from ubiquity to diversity. Trends Microbiol https://doi.org/10.1016/j.tim.2020.06.010.

254. Lestrate P, Delrue RM, Danese I, Didembourg C, Taminiau B, Mertens P, De Bolle X, Tibor A, Tang CM, Letesson JJ. 2000. Identification and characterization of in vivo attenuated mutants of Brucella melitensis. Mol Microbiol 38:543–551.

255. Kumar S, Kumar N, Alam N, Gourinath S. 2014. Crystal structure of serine acetyl transferase from Brucella abortus and its complex with coenzyme A. Biochim Biophys Acta - Proteins Proteomics 1844:1741–1748.

256. Rode LJ, Lankford CE, Schuhardt VT. 1951. Studies of sulfur metabolism of Brucella suis. J Bacteriol 62:571–582.

257. Devi S, Tarique KF, Ali MF, Abdul Rehman SA, Gourinath S. 2019. Identification and characterization of Helicobacter pylori O-acetylserine-dependent cystathionine β-synthase, a distinct member of the PLP-II family. Mol Microbiol 112:718–739.

258. Campanini B, Benoni R, Bettati S, Beck CM, Hayes CS, Mozzarelli A. 2015. Moonlighting O-acetylserine sulfhydrylase: New functions for an old protein. Biochim Biophys Acta - Proteins Proteomics 1854:1184–1193.

259. Dharavath S, Raj I, Gourinath S. 2017. Structure-based mutational studies of O-acetylserine sulfhydrylase reveal the reason for the loss of cysteine synthase complex formation in Brucella abortus. Biochem J 474:1221–1239.

260. Korshunov S, Imlay KRC, Imlay JA. 2020. Cystine import is a valuable but risky process whose hazards Escherichia coli minimizes by inducing a cysteine exporter. Mol Microbiol 113:22–39.

261. Lambeth JD. 2004. NOX enzymes and the biology of reactive oxygen. Nat Rev Immunol 4:181–189.

262. Sternon JF, Godessart P, de Freitas RG, Van der Henst M, Poncin K, Francis N, Willemart K, Christen M, Christen B, Letesson JJ, De Bolle X. 2018. Transposon sequencing of Brucella abortus uncovers essential genes for growth in vitro and inside macrophages. Infect Immun 86:1–20.

263. Khan SR, Gaines J, Roop RM, Farrand SK. 2008. Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. Appl Environ Microbiol 74:5053–5062.

264. García Lobo JM, Ortiz Y, Gonzalez-Riancho C, Seoane A, Arellano-Reynoso B, Sangari FJ. 2019. Polymorphisms in Brucella Carbonic Anhydrase II Mediate CO2 Dependence and Fitness in vivo. Front Microbiol 10:1–13.

265. Smith EP, Cotto-Rosario A, Borghesan E, Held K, Miller CN, Celli J. 2020. Epistatic Interplay between Type IV Secretion Effectors Engages the Small GTPase Rab2 in the Brucella Intracellular Cycle. MBio 11:e03350-19.

266. de Barsy M, Jamet A, Filopon D, Nicolas C, Laloux G, Rual JF, Muller A, Twizere JC, Nkengfac B, Vandenhaute J, Hill DE, Salcedo SP, Gorvel JP, Letesson JJ, De Bolle X. 2011. Identification of a Brucella spp. secreted effector specifically interacting with human small GTPase Rab2. Cell Microbiol 13:1044–1058.

267. Toh-e A, Ohkusu M, Shimizu K, Ishiwada N, Watanabe A, Kamei K. 2018. Novel biosynthetic pathway for sulfur amino acids in Cryptococcus neoformans. Curr Genet 64:681–696.

268. Ferber DM, Ely B. 1982. Resistance to amino acid inhibition in Caulobacter crescentus. Mol Gen Genet MGG 187:446–452.

269. Hullo M-F, Auger S, Soutourina O, Barzu O, Yvon M, Danchin A, Martin-Verstraete I. 2007. Conversion of methionine to cysteine in Bacillus subtilis and its regulation. J Bacteriol 189:187–197.

270. Devi S, Abdul Rehman SA, Tarique KF, Gourinath S. 2017. Structural characterization and functional analysis of cystathionine β-synthase: an enzyme involved in the reverse transsulfuration pathway of Bacillus anthracis. FEBS J 284:3862–3880.

271. Matoba Y, Yoshida T, Izuhara-Kihara H, Noda M, Sugiyama M. 2017. Crystallographic and mutational analyses of cystathionine β-synthase in the H(2) S-synthetic gene cluster in Lactobacillus plantarum. Protein Sci 26:763–783.

272. Purcell EB, Crosson S. 2008. Photoregulation in prokaryotes. Curr Opin Microbiol2008/04/11. 11:168–178.

273. Losi A, Polverini E, Quest B, Gartner W. 2002. First evidence for phototropin-related blue-light receptors in prokaryotes. Biophys J2002/04/20. 82:2627–2634.

274. Christie JM, Salomon M, Nozue K, Wada M, Briggs WR. 1999. LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide. Proc Natl Acad Sci U S A1999/07/21. 96:8779–8783.

275. Glantz ST, Carpenter EJ, Melkonian M, Gardner KH, Boyden ES, Wong GK, Chow BY. 2016. Functional and topological diversity of LOV domain photoreceptors. Proc Natl Acad Sci U S A 113:E1442-51.

276. Herrou J, Crosson S. 2011. Function, structure and mechanism of bacterial photosensory LOV proteins. Nat Rev Microbiol 9:713–723.

277. Losi A, Mandalari C, Gartner W. 2015. The Evolution and Functional Role of Flavin-based Prokaryotic Photoreceptors. Photochem Photobiol2015/07/04. 91:1021–1031.

278. Villar E, Vannier T, Vernette C, Lescot M, Cuenca M, Alexandre A, Bachelerie P, Rosnet T, Pelletier E, Sunagawa S, Hingamp P. 2018. The Ocean Gene Atlas: exploring the biogeography of plankton genes online. Nucleic Acids Res2018/05/23. 46:W289–W295.

279. Batut J, Andersson SGE, O'Callaghan D. 2004. The evolution of chronic infection strategies in the α -proteobacteria. Nat Rev Microbiol 2:933–945.

280. Herrou J, Crosson S, Fiebig A. 2017. Structure and function of HWE/HisKA2-family sensor histidine kinases. Curr Opin Microbiol 36:47–54.

281. Hoch JA, Silhavy TJ. 1995. Two-Component Signal Transduction. ASM Press, Washington, D.C.

282. Purcell EB, Siegal-Gaskins D, Rawling DC, Fiebig A, Crosson S. 2007. A photosensory two-component system regulates bacterial cell attachment. Proc Natl Acad Sci U S A 104:18241–18246.

283. Bonomi HR, Posadas DM, Paris G, Carrica Mdel C, Frederickson M, Pietrasanta LI, Bogomolni RA, Zorreguieta A, Goldbaum FA. 2012. Light regulates attachment, exopolysaccharide production, and nodulation in Rhizobium leguminosarum through a LOV-histidine kinase photoreceptor. Proc Natl Acad Sci U S A 109:12135–12140.

284. Fiebig A, Herrou J, Fumeaux C, Radhakrishnan SK, Viollier PH, Crosson S. 2014. A cell cycle and nutritional checkpoint controlling bacterial surface adhesion. PLoS Genet 10:e1004101.

285. Reyes Ruiz LM, Fiebig A, Crosson S. 2019. Regulation of bacterial surface attachment by a network of sensory transduction proteins. PLoS Genet 15:e1008022.

286. Foreman R, Fiebig A, Crosson S. 2012. The LovK-LovR two-component system is a regulator of the general stress pathway in Caulobacter crescentus. J Bacteriol 194:3038–3049.

287. Kaczmarczyk A, Hochstrasser R, Vorholt JA, Francez-Charlot A. 2014. Complex two-component signaling regulates the general stress response in Alphaproteobacteria. Proc Natl Acad Sci U S A 111:E5196-204.

288. Gottschlich L, Bortfeld-Miller M, Gabelein C, Dintner S, Vorholt JA. 2018. Phosphorelay through the bifunctional phosphotransferase PhyT controls the general stress response in an alphaproteobacterium. PLoS Genet 14:e1007294.

289. Lori C, Kaczmarczyk A, de Jong I, Jenal U. 2018. A Single-Domain Response Regulator Functions as an Integrating Hub To Coordinate General Stress Response and Development in Alphaproteobacteria. MBio2018/05/24. 9.

290. Correa F, Ko WH, Ocasio V, Bogomolni RA, Gardner KH. 2013. Blue light regulated two-component systems: enzymatic and functional analyses of light-oxygen-voltage (LOV)-histidine kinases and downstream response regulators. Biochemistry 52:4656–4666.

291. Francez-Charlot A, Frunzke J, Reichen C, Ebneter JZ, Gourion B, Vorholt JA. 2009. Sigma factor mimicry involved in regulation of general stress response. Proc Natl Acad Sci U S A2009/02/17. 106:3467–3472.

292. Sycz G, Carrica MC, Tseng TS, Bogomolni RA, Briggs WR, Goldbaum FA, Paris G. 2015. LOV Histidine Kinase Modulates the General Stress Response System and Affects the virB Operon Expression in Brucella abortus. PLoS One2015/05/21. 10:e0124058.

293. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, Djahanschiri B, Zeller G, Mende DR, Alberti A, Cornejo-Castillo FM, Costea PI, Cruaud C, d'Ovidio F, Engelen S, Ferrera I, Gasol JM, Guidi L, Hildebrand F, Kokoszka F, Lepoivre C, Lima-Mendez G, Poulain J, Poulos BT, Royo-Llonch M, Sarmento H, Vieira-Silva S, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Tara Oceans coordinators, Bowler C, de Vargas C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Jaillon O, Not F, Ogata H, Pesant S, Speich S, Stemmann L, Sullivan MB, Weissenbach J, Wincker P, Karsenti E, Raes J, Acinas SG, Bork P. 2015. Structure and function of the global ocean microbiome. Science (80-)2015/05/23. 348:1261359.

294. Tully BJ, Graham ED, Heidelberg JF. 2018. The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. Sci Data2018/01/18. 5:170203.

295. Koblizek M, Masin M, Ras J, Poulton AJ, Prasil O. 2007. Rapid growth rates of aerobic anoxygenic phototrophs in the ocean. Env Microbiol2007/09/07. 9:2401–2406.

296. Kolber ZS, Plumley FG, Lang AS, Beatty JT, Blankenship RE, VanDover CL, Vetriani C, Koblizek M, Rathgeber C, Falkowski PG. 2001. Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean. Science (80-)2001/06/30. 292:2492–2495.

297. Oh HM, Giovannoni SJ, Ferriera S, Johnson J, Cho JC. 2009. Complete genome sequence of Erythrobacter litoralis HTCC2594. J Bacteriol 191:2419–2420.

298. Wang Y, Zhang R, Zheng Q, Jiao N. 2014. Draft Genome Sequences of Two Marine Phototrophic Bacteria, Erythrobacter longus Strain DSM 6997 and Erythrobacter litoralis Strain DSM 8509. Genome Announc 2.

299. Delmont TO, Quince C, Shaiber A, Esen OC, Lee ST, Rappe MS, McLellan SL, Lucker S, Eren AM. 2018. Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. Nat Microbiol2018/06/13. 3:804–813.

300. Lei X, Zhang H, Chen Y, Li Y, Chen Z, Lai Q, Zhang J, Zheng W, Xu H, Zheng T. 2015. Erythrobacter luteus sp. nov., isolated from mangrove sediment. Int J Syst Evol Microbiol2015/04/26. 65:2472–2478.

301. Yurkov V, Stackebrandt E, Holmes A, Fuerst JA, Hugenholtz P, Golecki J, Gad'on N, Gorlenko VM, Kompantseva EI, Drews G. 1994. Phylogenetic positions of novel aerobic, bacteriochlorophyll a-containing bacteria and description of Roseococcus thiosulfatophilus gen. nov., sp. nov., Erythromicrobium ramosum gen. nov., sp. nov., and Erythrobacter litoralis sp. nov. Int J Syst Bacteriol 44:427–434.

302. Yurkov V, Van Gemerden H. 1993. Abundance and salt tolerance of obligately aerobic, phototrophic bacteria in a marine microbial mat. Netherlands J Sea Res 31:57–62.

303. Dikiy I, Edupuganti UR, Abzalimov RR, Borbat PP, Srivastava M, Freed JH, Gardner KH. 2019. Insights into histidine kinase activation mechanisms from the monomeric blue light sensor EL346. Proc Natl Acad Sci U S A https://doi.org/10.1073/pnas.1813586116.

304. Rivera-Cancel G, Ko WH, Tomchick DR, Correa F, Gardner KH. 2014. Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. Proc Natl Acad Sci U S A 111:17839–17844.

305. Zheng Q, Lin W, Liu Y, Chen C, Jiao N. 2016. A Comparison of 14 Erythrobacter Genomes Provides Insights into the Genomic Divergence and Scattered Distribution of Phototrophs. Front Microbiol 7:984.

306. Jiang XW, Cheng H, Huo YY, Xu L, Wu YH, Liu WH, Tao FF, Cui XJ, Zheng BW. 2018. Biochemical and genetic characterization of a novel metallo-beta-lactamase from marine bacterium Erythrobacter litoralis HTCC 2594. Sci Rep2018/01/18. 8:803.

307. Ulrich LE, Zhulin IB. 2010. The MiST2 database: a comprehensive genomics resource on microbial signal transduction. Nucleic Acids Res2009/11/11. 38:D401-7.

308. Marks ME, Castro-Rojas CM, Teiling C, Du L, Kapatral V, Walunas TL, Crosson S. 2010. The genetic basis of laboratory adaptation in Caulobacter crescentus. J Bacteriol2010/05/18. 192:3678–3688.

309. Crosson S, Moffat K. 2001. Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction. Proc Natl Acad Sci U S A2001/03/15. 98:2995–3000.

310. Salomon M, Christie JM, Knieb E, Lempert U, Briggs WR. 2000. Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. Biochemistry2000/08/05. 39:9401–9410.

311. Crosson S, Moffat K. 2002. Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch. Plant Cell2002/05/30. 14:1067–1075.

312. Salomon M, Eisenreich W, Durr H, Schleicher E, Knieb E, Massey V, Rudiger W, Muller F, Bacher A, Richter G. 2001. An optomechanical transducer in the blue light receptor phototropin from Avena sativa. Proc Natl Acad Sci U S A2001/10/19. 98:12357–12361.

313. Swartz TE, Corchnoy SB, Christie JM, Lewis JW, Szundi I, Briggs WR, Bogomolni RA. 2001. The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. J Biol Chem2001/07/10. 276:36493–36500.

314. Staron A, Sofia HJ, Dietrich S, Ulrich LE, Liesegang H, Mascher T. 2009. The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. Mol Microbiol2009/09/10. 74:557–581.

315. Martinez-Salazar JM, Salazar E, Encarnacion S, Ramirez-Romero MA, Rivera J. 2009. Role of the extracytoplasmic function sigma factor RpoE4 in oxidative and osmotic stress responses in Rhizobium etli. J Bacteriol2009/04/21. 191:4122–4132.

316. Sauviac L, Philippe H, Phok K, Bruand C. 2007. An extracytoplasmic function sigma factor acts as a general stress response regulator in Sinorhizobium meliloti. J Bacteriol2007/04/03. 189:4204–4216.

317. Alvarez-Martinez CE, Lourenco RF, Baldini RL, Laub MT, Gomes SL. 2007. The ECF sigma factor sigma(T) is involved in osmotic and oxidative stress responses in Caulobacter crescentus. Mol Microbiol2007/11/08. 66:1240–1255.

318. Gourion B, Francez-Charlot A, Vorholt JA. 2008. PhyR is involved in the general stress response of Methylobacterium extorquens AM1. J Bacteriol2007/11/21. 190:1027–1035.

319. Gourion B, Sulser S, Frunzke J, Francez-Charlot A, Stiefel P, Pessi G, Vorholt JA, Fischer HM. 2009. The PhyR-sigma(EcfG) signalling cascade is involved in stress response and symbiotic efficiency in Bradyrhizobium japonicum. Mol Microbiol2009/06/27. 73:291–305.

320. Jans A, Vercruysse M, Gao S, Engelen K, Lambrichts I, Fauvart M, Michiels J. 2013. Canonical and non-canonical EcfG sigma factors control the general stress response in Rhizobium etli. Microbiologyopen2013/12/07. 2:976–987.

321. Lourenco RF, Kohler C, Gomes SL. 2011. A two-component system, an anti-sigma factor and two paralogous ECF sigma factors are involved in the control of general stress response in Caulobacter crescentus. Mol Microbiol2011/05/14. 80:1598–1612.

322. Ginter C, Kiburu I, Boudker O. 2013. Chemical catalysis by the translocator protein (18 kDa). Biochemistry2013/05/09. 52:3609–3611.

323. Yeliseev AA, Kaplan S. 1995. A sensory transducer homologous to the mammalian peripheral-type benzodiazepine receptor regulates photosynthetic membrane complex formation in Rhodobacter sphaeroides 2.4.1. J Biol Chem1995/09/08. 270:21167–21175.

324. Fuller KK, Ringelberg CS, Loros JJ, Dunlap JC. 2013. The fungal pathogen Aspergillus fumigatus regulates growth, metabolism, and stress resistance in response to light. MBio2013/03/28. 4.

325. Zhu YS, Hearst JE. 1986. Regulation of expression of genes for light-harvesting antenna proteins LH-I and LH-II; reaction center polypeptides RC-L, RC-M, and RC-H; and enzymes of bacteriochlorophyll and carotenoid biosynthesis in Rhodobacter capsulatus by light and oxygen. Proc Natl Acad Sci U S A1986/10/01. 83:7613–7617.

326. Tomasch J, Gohl R, Bunk B, Diez MS, Wagner-Dobler I. 2011. Transcriptional response of the photoheterotrophic marine bacterium Dinoroseobacter shibae to changing light regimes. ISME J2011/06/10. 5:1957–1968.

327. Staroń A, Mascher T. 2010. General stress response in a-proteobacteria: PhyR and beyond. Mol Microbiol 78:271–277.

328. Hengge R. 2008. The two-component network and the general stress sigma factor RpoS (sigma S) in Escherichia coli. Adv Exp Med Biol2008/09/17. 631:40–53.

329. Akbar S, Gaidenko TA, Kang CM, O'Reilly M, Devine KM, Price CW. 2001. New family of regulators in the environmental signaling pathway which activates the general stress transcription factor sigma(B) of Bacillus subtilis. J Bacteriol2001/02/07. 183:1329–1338.

330. Avila-Perez M, Hellingwerf KJ, Kort R. 2006. Blue light activates the sigmaB-dependent stress response of Bacillus subtilis via YtvA. J Bacteriol2006/08/23. 188:6411–6414.

331. O'Donoghue B, NicAogain K, Bennett C, Conneely A, Tiensuu T, Johansson J, O'Byrne C. 2016. Blue-Light Inhibition of Listeria monocytogenes Growth Is Mediated by Reactive Oxygen Species and Is Influenced by sigmaB and the Blue-Light Sensor Lmo0799. Appl Env Microbiol2016/05/01. 82:4017–4027.

332. Ondrusch N, Kreft J. 2011. Blue and red light modulates SigB-dependent gene transcription, swimming motility and invasiveness in Listeria monocytogenes. PLoS One2011/01/26. 6:e16151.

333. Moriconi V, Sellaro R, Ayub N, Soto G, Rugnone M, Shah R, Pathak GP, Gartner W, Casal JJ. 2013. LOV-domain photoreceptor, encoded in a genomic island, attenuates the virulence of Pseudomonas syringae in light-exposed Arabidopsis leaves. Plant J2013/07/20. 76:322–331.

334. Braatsch S, Moskvin O V, Klug G, Gomelsky M. 2004. Responses of the Rhodobacter sphaeroides transcriptome to blue light under semiaerobic conditions. J Bacteriol 186:7726–7735.

335. Ziegelhoffer EC, Donohue TJ. 2009. Bacterial responses to photo-oxidative stress. Nat Rev Microbiol 7:856–863.

336. Swartz TE, Tseng TS, Frederickson MA, Paris G, Comerci DJ, Rajashekara G, Kim JG, Mudgett MB, Splitter GA, Ugalde RA, Goldbaum FA, Briggs WR, Bogomolni RA. 2007. Blue-light-activated histidine kinases: two-component sensors in bacteria. Science (80-)

317:1090–1093.

337. Endres S, Granzin J, Circolone F, Stadler A, Krauss U, Drepper T, Svensson V, Knieps-Grunhagen E, Wirtz A, Cousin A, Tielen P, Willbold D, Jaeger KE, Batra-Safferling R. 2015. Structure and function of a short LOV protein from the marine phototrophic bacterium Dinoroseobacter shibae. BMC Microbiol 15:30.

338. Finan TM, Kunkel B, De Vos GF, Signer ER. 1986. Second symbiotic megaplasmid in Rhizobium meliloti carrying exopolysaccharide and thiamine synthesis genes. J Bacteriol 167:66–72.

339. Hmelo LR, Borlee BR, Almblad H, Love ME, Randall TE, Tseng BS, Lin C, Irie Y, Storek KM, Yang JJ, Siehnel RJ, Howell PL, Singh PK, Tolker-Nielsen T, Parsek MR, Schweizer HP, Harrison JJ. 2015. Precision-engineering the Pseudomonas aeruginosa genome with two-step allelic exchange. Nat Protoc 10:1820–1841.

340. Ried JL, Collmer A. 1987. An nptI-sacB-sacR cartridge for constructing directed, unmarked mutations in gram-negative bacteria by marker exchange-eviction mutagenesis. Gene 57:239–246.

341. Pitcher DG, Saunders NA, Owen RJ. 1989. Rapid Extraction of Bacterial Genomic DNA with Guanidium Thiocyanate. Lett Appl Microbiol 8:151–156.

342. Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen IMA, Pati A, Nielsen T, Markowitz VM, Kyrpides NC. 2015. The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4). Stand Genomic Sci 10:86.

343. Shiba T, Simidu U. 1982. Erythrobacter longus gen. nov., sp. nov., an aerobic bacterium which contains Bacteriochlorophyll a. Int J Syst Evol Microbiol 32:211–217.

344. Francis CA, Co EM, Tebo BM. 2001. Enzymatic manganese(II) oxidation by a marine alpha-proteobacterium. Appl Env Microbiol 67:4024–4029.

345. Koblizek M, Beja O, Bidigare RR, Christensen S, Benitez-Nelson B, Vetriani C, Kolber MK, Falkowski PG, Kolber ZS. 2003. Isolation and characterization of Erythrobacter sp. strains from the upper ocean. Arch Microbiol 180:327–338.

346. Koblizek M, Janouskovec J, Obornik M, Johnson JH, Ferriera S, Falkowski PG. 2011. Genome sequence of the marine photoheterotrophic bacterium Erythrobacter sp. strain NAP1. J Bacteriol 193:5881–5882.