Title

Genetics and Genomics of Perchlorate Reduction and Reactive Chlorine Species

Author

Melnyk, Ryan Alexander

Publication Date

2014

Peer reviewed|Thesis/dissertation

Genetics and Genomics of Perchlorate Reduction and Reactive Chlorine Species

By

Ryan Alexander Melnyk

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John D. Coates, Chair
Professor Kathleen R. Ryan
Professor Peter H. Quail

Fall 2014

Genetics and Genomics of Perchlorate Reduction and Reactive Chlorine Species

by Ryan Alexander Melnyk

Abstract

Genetics and Genomics of Perchlorate Reduction and Reactive Chlorine Species

by

Ryan Alexander Melnyk

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor John D. Coates, Chair

The anion perchlorate (ClO_4^-) is a soluble compound that accumulates very slowly in the natural environment as the result of an unknown mechanism. Owing to its excellent oxidizing capability, large quantities of various perchlorate salts are also generated anthropogenically as components of various types of explosives and rocket fuel. Unregulated disposal of perchlorate as well as accidental release has led to elevated amounts of perchlorate in drinking water sources, where it can interfere with thyroid function. Due to its adverse health effects, the United States Environmental Protection Agency has elected to regulate perchlorate under the Safe Water Drinking Act as of 2012.

The high solubility of perchlorate and its strong oxidizing potential make it an attractive terminal electron acceptor for bacteria, and unsurprisingly, there are bacteria that utilize it in such a manner. This metabolism depends on two key enzymes that have been studied extensively: perchlorate reductase (PcrA) and chlorite dismutase (Cld). Cloning and sequencing of these genes revealed that they are part of the same operon, and preliminary phylogenetic analysis indicated that the histories of these genes were not congruent with phylogeny based on the 16S genes of these organisms. These observations, coupled with the fact that perchlorate reducers are found in disparate taxonomic groups, led to the hypothesis that the genetic constituents of perchlorate reduction are horizontally transferred.

The acquisition of several complete (*Azospira suillum* PS and *Dechloromonas aromatica* RCB) and draft genomes (*Magnetospirillum bellicus* VDY and *Dechloromonas agitata* CKB) of dissimilatory perchlorate reducing bacteria (DPRB) were sequenced in order to understand more about the genomic basis of this metabolism. Another strain that does not reduce perchlorate, *Dechloromonas* sp. JJ, was also sequenced and used a baseline for comparisons. This led to the identification of a genomic island centered

around the *pcr* and *cld* genes, but in addition contained many conserved genes of unknown function, with possible roles in electron transport and regulation (Chapter 1). This genomic island was designated the perchlorate reduction genomic island (PRI).

To understand the role, if any, these genes in the PRI play in perchlorate reduction metabolism, a tractable genetic system with methods for mutagenesis and complementation was needed. To fill this role, the organism *Azospira suillum* PS was developed into a model system for genetic manipulations (Chapter 2). Physiological characterizations were also performed to establish a baseline phenotype with respect to growth on nitrate, chlorate, and perchlorate, as well as combinations thereof. A rich media for rapid and routine aerobic culturing of PS was also developed.

Using the genetic and physiological tools developed, a library of PS transposon mutants were generated and screened for defects specifically during growth on perchlorate. Many of these transposon insertions were found to be in the genes in the PRI, so individual deletions of all 17 genes were made. Eight of these genes were essential for perchlorate reduction, and complementation of each of the eight deletions restored the phenotype. These genetic methods were instrumental in defining the set of essential genes for perchlorate reduction in PS. (Chapter 2).

Several of the genes in the PRI were not essential for perchlorate reduction including homologs of the *sigF-nrsF* sigma factor/anti-sigma factor system. In PS, this system is responsive to chlorite stress, and governs transcription of an adjacent operon in the PRI (Chapter 3). The genes in this operon consist of a small periplasmic methionine-rich protein (*mrpX*) and a putative methionine sulfoxide reductase (*yedYZ*). This system responds to the toxic molecule hypochlorite produced as a side product during perchlorate reduction, which would be specifically scavenged in the periplasm by the MrpX protein and regenerated by YedYZ. A comparative genomic analysis indicates that this mechanism is broadly conserved in the Proteobacteria, and tends to be enriched in organisms with known associations with eukaryotic hosts, both plant and animal.

Further sequencing of DPRB isolates has led to the accumulation of a large library of complete and draft genomes, many of which are from the same genus or species. A comprehensive phylogenomic analysis of these organisms revealed the full extent to which horizontal transfer of the PRI has shaped this metabolism (Chapter 4). In two separate instances, the PRI of divergent members of the same species exhibits almost complete nucleotide identity, indicating that the PRI is under both strong positive selection and very frequent horizontal gene transfer. Additionally, several of the accessory genes in the PRI are polyphyletic and individual lineages can be traced to genes in the chromosomes of phylogenetically similar organisms.

The work presented in this dissertation represents an important transition in the study of the molecular biology of perchlorate reduction. The identification of the PRI

provided a plausible mechanism for how this metabolism is horizontally inherited. The genetic tools developed for *Azospira suillum* PS led to the discovery of many novel genes and also confirmed the important role that reactive chlorine species play in this metabolism. Finally, the large library of DPRB genomes and comparative analysis decribed here provide a solid base as perchlorate reduction research increasingly moves towards community analysis and *in situ* function.

"Every book is a quotation…and every man is a quotation from all his ancestors."

-Ralph Waldo Emerson

For my grandparents

Ruth E. Frey, Roman A. Melnyk, and Lubomyra M. Melnyk

I know you would be proud

And their children, whose love and wisdom guides my past, present, and future

Adrian L. Melnyk and Sue M. Melnyk

Table of Contents

List of Figures

List of Tables

Acknowledgements

Any list of people I need to acknowledge for helping me make it this far down this path will be woefully incomplete. As a baseball enthusiast, I know that I'm crossing home plate today not because I hit a leadoff home run, but because I've been at the top of a great batting order. None of this would have happened without a family who cared about my education, access to public schools with teachers that cared, and financial support for a high-quality post-secondary education. I realize that I'm lucky (or blessed, or privileged, depending on your perspective) and would encourage whoever is reading this to understand that as far as science education goes in America, there is an ever-widening chasm separating the extremes of quality, and those of us in science must do what we can to bridge that gap. I envision a future where every student, no matter how rich or how poor, with the aptitude and talent finds that little spark, that glimmer of understanding that can lead to the love of science that so many of us have here at Berkeley. Right now, we are long way from that future.

Alright, soapbox aside. My time here at Berkeley in PMB has been shaped by some unique individuals who deserve mentioning here. To Dr. Alex Schultink and Thomas Kleist and the rest of my cohort, thank you for making the transition to grad school a fun one. To all of those who participated in my fantasy football and baseball leagues, thank you for the trash talk, photoshop wars, GIF threads, and welcome distraction from science. If it weren't for you guys, I might have graduated in four years (albeit with fewer smiles)! To Adam Steinbrenner and David Hershey, my fellow Phillies-turned-A's fans, I never thought that I would find such close friends in grad school, let alone on the 2nd floor of Koshland. Thank you for countless trips to Bobby G's, pitchers of beer, hikes, A's games, and discussions of baseball and science. It's these things that I will remember long after we all leave grad school. To Robbie Calderon, Stephen Bird, and Mikel Shybut, I couldn't ask for better friends or roommates. It's always fun to come home and listen to your unique perspective on science and life. To the rest of the PMB softball team, The Foul Tips, I raise my glass (of Fireball) to you and urge you to keep it **FOUL** and keep fighting against the up-tight MCB teams that would restrain our Tips…who am I kidding though, I'm not going to stop coming to games next year. To eveyone else in the PMB department, thank you for making it such a social, fun, and close-knit atmosphere: this is a great place to be a grad student and you folks are a main reason why.

I'd like to thank Dr. Arash Komeili and Dr. Dorothée Murat for teaching a freshly graduated "rotato" who knew nothing about molecular biology all of the skills needed to get a genetics project started from scratch. Arash, thank you also for being a "hands-on" Quals chair and helping me develop the early stages of my project along with the rest of my committee (Dr. Michi Taga, Dr. Cheryl Kerfeld, and Dr. Chris Anderson). Dr.

Kathleen Ryan and Dr. Peter Quail, thank you for serving on my dissertation committee and your insight and earnest curiosity about my project.

For the last five years, I have been truly excited about coming into the Coates Lab each day thanks to an array of talented co-workers who are a joy to be around. First out of this crew is Dr. Kelly Wrighton, who trained me in the Coates Lab but treated me as an equal collaborator and encouraged me to share my ideas about her project from day one and continued to provide feedback on my project during her postdoc. Dr. Cameron Thrash deserves a mention as well for teaching me to make an awesome bloody mary and giving me a lot of advice on phylogenomics. You both provided excellent role models of how to be driven, independent scientists who work hard (and play hard). I am so proud of you both as you go about launching your own labs and I can't wait to see where your research takes you.

Our lab manager, Anna Engelbrektson, I can't thank you enough for helping a few brand-new grad students get organized and getting the Coates Lab running smoothly after a big personnel turnover. Although it may seem that you crushed the rebellion in Backlabistan, the people will rise again! Charlotte Carlström, it's been awesome to see how we have developed as scientists since we were first-years together. I don't think I ever would have guessed that our projects would have converged like they are now. Israel Figueroa, the **Pun**isher, thank you for always being a source of levity and humor in the lab. Dr. Matt Youngblut and Dr. Misha Mehta, it has been awesome working with both of you over the last year or so. Each of you has brought a new dimension to the lab both professionally and socially. Matt, keep sending everyone who dislikes football (the American kind) into the cold room Monday mornings. Misha, I hope you are working on memorizing locus tags because I need you to take over as PS Rain Woman. Ouwei Wang, I'm so glad that you are charging ahead with so many of the new questions we have about perchlorate reduction. I can't wait to see where you take this project. Pravin Shrestha, I salute you for doing the literally dirty job of getting these columns up and running. Dana Loutey and Maggie Stoeva, it's been great to see such talented young female scientists come into the lab to infuse new energy into projects and I wish you both continued success.

I could not have finished my PhD if it weren't for two individuals that I've spent the last five years discussing, arguing, eating lunch, and laughing nearly every day with me. Dr. Hans Carlson, I have to thank you for coming into the Coates Lab and helping me with the godforsaken JR project. Your creative and collaborative approach to science is something I try to emulate. Thank you for guiding me through a successful quals proposal and encouraging me to branch out into genetics and biochemistry in the Coates Lab. Iain Clark, we made it! My Backlabistan comrade, I can't count how many times your rational approach to experimental design and science has saved me from doing something stupid or helped me work through some seemingly insurmountable

obstacle (cough PS genetics cough cough). Thanks for sharing scripts, troubleshooting growth curves, and staring at incomprehensible data with me. Hans and Iain, thanks for being incredible role models and mentors, in realms professional and personal. You have both become husbands and recently fathers during the time I've known you, and I wish you and your budding families much continued happiness.

To John Coates, my mentor and friend for the last five years. I have learned so much from you scientifically, professionally, and personally and I owe you so much for your willingness to allow grad students to blaze their own trails. From my first day of my Coates Lab rotation, I felt like an equal whose opinion and insight mattered. This quality alone shaped my time here at Berkeley more than anything else, helping me to gain confidence in my abilities as an independent scientist. Of course, that's merely the first thing on a long list including such things as: "the mushroom treatment," the "suck it and see experiment," and of course "this shit's too easy." In all seriousness though, thank you for being an advocate for your students and all of your advice. I thank you so much for passing on your knowledge of microbial physiology and showing me that is possible to simultaneously be a successful faculty member, a fun boss who cares about his students and postdocs, and most importantly, an engaged and supportive father and husband.

To Whitney Clark: I know I've only known you for a little while, but it feels like a lot longer (and I mean that in a GOOD way!). Thank you for making my post-grad school decision to stay in the Bay Area easy, and thank you for always putting a smile on my face by being such a "cheerful little chirping bird!" I can't wait to go on more adventures with you; you bring a lightness into my life that sustains me in more ways than one. I love you.

Finally to my parents Sue and Adrian Melnyk, thank you for all of your support and encouraging me to seek out knowledge and happiness. Everything about the man I am today I owe in some way to you guys and how you both taught me to remember "who you are and whose you are". I love you both.

Chapter 1:

A comparative analysis of the genomes of four perchlorate reducers reveals a genomic island with novel regulatory and metabolic genes

Abstract

A comparative analysis of the genomes of four dissimilatory (per)chlorate reducing bacteria (DPRB) has revealed a genomic island associated with perchlorate reduction. In addition to the characterized metabolic genes for perchlorate reductase and chlorite dismutase, the perchlorate reduction genomic island (PRI) contains multiple conserved uncharacterized genes possibly involved in electron transport and regulation.

The previously finished genome from the organism *Dechloromonas aromatica* RCB was compared to the newly finished genome from *Azospira suillum* PS. Additionally, we generated draft genomes for the DPRB *Dechloromonas agitata* CKB and *Magnetospirillum bellicus* VDY, as well as the non-perchlorate reducing strain *Dechloromonas* sp. JJ to use as a comparison. This expands our genomic database of DPRB to three genera across two phylogenetic classes (Alphaproteobacteria and Betaproteobacteria) and gives some insight into other roles that the PRI plays beyond the reduction of perchlorate to chloride.

This chapter was originally published as a report in *Applied and Environmental Microbiology* (Melnyk *et al*., 2011. **77**:7401-7404).

Introduction

Perchlorate is a stable, soluble, and toxic anion that is deposited in the environment by both natural and anthropogenic means (Kounaves et al., 2010). It represents a potential health risk to humans, as it accumulates in food sources (Kirk et al., 2005; Sanchez et al., 2005), where it can interfere with thyroid function upon ingestion (Lawrence et al., 2000). As a result of its persistence in the environment, certain bacteria have evolved a mechanism to reduce perchlorate and chlorate [collectively (per)chlorate] coupled to energy metabolism (Coates and Achenbach, 2004). The two key enzymes in the perchlorate reduction pathway are perchlorate reductase and chlorite dismutase, encoded by the *pcrABCD* and *cld* genes, respectively (van Ginkel et al., 1996; Bender et al., 2002; 2005).

The first (per)chlorate-reducing organism to be genome sequenced was *Dechloromonas aromatica* RCB (Coates et al., 2001). Recently, we have acquired genome sequences for several other perchlorate reducers including the finished genome of *Azospira suillum* strain PS (Achenbach et al., 2001), *Magnetospirillum bellicus* VDY (Thrash et al., 2010), and *D. agitata* CKB (Bruce et al., 1999; Achenbach et al., 2001) (DOE Joint Genome Institute and Eureka Genomics, CA). Additionally, the only known non-(per)chlorate-reducing member of the *Dechloromonas* genus, *Dechloromonas* sp. JJ (Coates et al., 2001), was also sequenced as part of this study (Eureka Genomics).

Identification of genes conserved in DPRB

The genera *Dechloromonas* and *Azospira* are both located within the *Rhodocyclaceae* family of the Betaproteobacteria, from which perchlorate-reducing organisms are frequently isolated (Coates et al., 1999). In contrast, *M. bellicus* VDY is a member of the Alphaproteobacteria, and shares 96% 16S rRNA gene sequence identity with the magnetotactic species *M. magnetotacticum* and *M. gryphiswaldense*, despite its apparent inability to form magnetosomes (Thrash et al., 2010) (Figure 1).

In our initial analysis of the *A. suillum* genome, we identified 11 genes adjacent to the genes encoding both perchlorate reductase and chlorite dismutase that were also present in *D. aromatica*, which was unsurprising, given their close phylogenetic relationship (Figure 1). To examine the complete extent of the synteny surrounding *pcr* and *cld* across all four perchlorate reducers, we used a simple protein similarity search (phmmer) (Eddy, 1998) using the translations of the genes surrounding *pcrA* from *D. aromatica* as a query against a database of 1109 curated and finished genomes from HAMAP (Lima et al., 2009) in addition to the draft genomes of *A. suillum, M. bellicus, D. agitata,* and *Dechloromonas* sp. JJ. We found 17 pairs of similar genes with conserved synteny between *A. suillum* and *D. aromatica*, in addition to a subset of these genes in *D. agitata* (10 genes) and *M. bellicus* (12 genes) (Figure 2). While in most cases the best-scoring hits from the *D. aromatica* queries were from the other (per)chlorate reducers,

there were a few aberrant instances. For example, the gene encoding the molybdopterin biosynthesis protein A (*moaA*) homolog in *M. bellicus* had only a moderate degree of similarity to the closely related *moaA* genes of *D. aromatica, A. suillum*, and *D. agitata*, despite its similar genomic location (Figure 2). *Dechloromonas* strain JJ did not have high-scoring hits to most of the 17 genes shared by *A. suillum* and *D. aromatica*, despite its close phylogenetic affiliation, and any homologs were non-contiguous, suggesting that there is no comparable genomic region in that organism.

Defining a genomic island associated with perchlorate reduction

While it has been proposed that *cld* has a phylogenetic history indicative of lateral gene transfer (Bender et al., 2004), the nature and extent of such a transfer event was unknown. However, the conservation of ten or more genes among four (per)chlorate reducers leads us to propose that this 10- to 25-kb region constitutes the core of a horizontally transferred (per)chlorate-reduction associated genomic island (PRI).

We justify the definition of this region as a genomic island based on the criteria presented by Juhas *et al.*, some of which we will summarize here (Juhas et al., 2009). Like most genomic islands, the GC content of the PRI core is different from the background GC content of the surrounding chromosome (Table 1). Additionally, the region that we have delineated as the PRI core in *D. aromatica* was identified as a genomic island by the web-based IslandViewer software, which integrates multiple methods for identifying genomic islands[IslandViewer citation goes here]. While we have not identified any flanking direct repeats indicative of site-specific integration or homologous recombination, there is a Pro tRNA gene within 3 kb of the PRI in *A. suillum* and 35 kb of the PRI core in *D. aromatica*. This was previously identified as a chromosomal integration site for a phage of *Rhizobium meliloti* (Semsey et al., 2002). In the genome of *M. bellicus*, there is no Pro tRNA in the vicinity of the PRI core, but a Glu tRNA is located 7 kb upstream of the *moaA* gene of the PRI. In the ~30 kb region downstream of the histidine kinase of the PRI in *M. bellicus*, there are two genes from the XerCD site-specific recombinase family, in addition to a restriction endonuclease/restriction site modification operon. The core of the PRI of *D. agitata* is located on a relatively small contig (19 kb), but we were still able to identify a transposase 2 kb away from the PRI core, which was not homologous to any gene from the PRIs of the other three genomes. In the ~35 kb region between the PRI core and the Pro tRNA in *D. aromatica,* there are homologs of the F plasmid conjugative transfer genes *traIDLEBVFWUNH* (Ippen-Ihler and Minkley, 1986). None of these gene families are present in the PRI of *A. suillum*, but there are multiple genes in its flanking regions annotated as "phage-associated". This genomic evidence suggests that the PRI may have been historically mobilized by conjugation, transformation, and transduction, and

integrated into host chromosomes via a variety of mechanisms. The fact that there is such heterogeneity among the non-core PRI genes of the four isolates, incuding three closely related strains from the *Rhodocyclaceae*, suggests that mobile elements containing the PRI core genes are extremely diverse and we have sampled only a fraction of that genetic diversity. It is worth noting that we have never observed a spontaneous mutant incapable of (per)chlorate reduction, unlike in other organisms where genomic island integration and excision can be directly observed (Fukuda et al., 2006; Domínguez et al., 2011). As a result, the exact boundaries delineating the PRI from the host chromosome may be ambiguous, particularly if there has been fixation of the PRI via erosion of site-specific excision loci or loss of a flanking direct repeat following a homologous recombination event.

Possible functions for the PRI beyond Pcr and Cld

The conservation of PRI genes outside of *pcrABCD* and *cld* suggests that there are additional regulatory and metabolic mechanisms involved in (per)chlorate reduction that are undescribed. The presence of the molybdenum cofactor biosynthesis protein homolog *moaA* is unsurprising, as perchlorate reduction is dependent on molybdenum, presumed to be due to the use of molybdopterin as a cofactor of PcrA (Bender et al., 2002; Chaudhuri et al., 2002; Bender et al., 2005). Additionally, there are three distinct families of *c*-type cytochromes in the PRI, none of which have an experimentally validated function, but we predict are involved in electron transport to perchlorate.

Furthermore, there are two separate modules in the PRI that we predict to regulate the transcription of (per)chlorate metabolism genes, including *pcrABCD* and *cld*. These two modules consist of a putative σ factor/anti-σ factor pair and a two-component system. The two-component system is of particular interest, as the response regulator contains a σ^{54} interaction/activation domain, and we have identified a promoter upstream of *pcrA* in *A. suillum* and *D. agitata* with binding sites for the σ^{54} factor RpoN and the regulator of anaerobic metabolism, FNR (Bender et al., 2005; Melnyk et al., 2011).

All perchlorate-reducing organisms utilize oxygen as an electron acceptor, and all but one (*D. agitata*) can sustain growth via denitrification. Both of these electron acceptors are preferred to perchlorate by model perchlorate reducers, including *D. aromatica* RCB and *A. suillum* PS. Because σ^{54}-interacting response regulators generally act as transcriptional activators, we propose that the two-component system is active in the absence of nitrate, and FNR is activated anaerobically. When both of these systems are active, transcription of *pcrABCD* would be induced. The absence of this regulatory system in *D. agitata* is supportive of such a hypothesis, as this organism is unable to

denitrify, and does not need to repress transcription of *pcrABCD* when nitrate is available.

Conclusion

This is the first identification of conserved gene families linked with the perchlorate reductase and chlorite dismutase genes in (per)chlorate-reducing organisms. The architecture of the PRI and surrounding chromosomal regions suggests that these genes have been co-transferred, and are completely absent in investigated organisms closely related to (per)chlorate reducers (e.g. *Dechloromonas* sp. JJ and *M. magnetotacticum*). These findings imply that the PRI represents a modular metabolism, in that it packages necessary regulatory elements with metabolic genes, in order to facilitate the integration of the horizontally transferred PRI into the host metabolism. To the best of our knowledge this represents only the second example of a genomic island that encodes a respiratory process in addition to the recently identified genomic island involved in organohalide reduction (McMurdie et al., 2009; 2011). However, the PRI is the first respiratory genomic island known that crosses broad phylogenetic boundaries (class, order, and family), as the organohalide genomic island has only been observed in the genus *Dehalococcoides*.

References

Abascal, F., Zardoya, R., and Posada, D. (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics (Oxford, England)* **21**: 2104–2105.

Achenbach, L., Michaelidou, U., Bruce, R., Fryman, J., and Coates, J. (2001) Dechloromonas agitata gen. nov., sp. nov. and Dechlorosoma suillum gen. nov., sp. nov., two novel environmentally dominant (per)chlorate-reducing bacteria and their phylogenetic position. *Int J Syst Evol Microbiol* **51**: 527–533.

Bender, K., Shang, C., Chakraborty, R., Belchik, S., Coates, J., and Achenbach, L. (2005) Identification, characterization, and classification of genes encoding perchlorate reductase. *J Bacteriol* **187**: 5090–5096.

Bender, K.S., O'Connor, S.M., Chakraborty, R., Coates, J.D., and Achenbach, L.A. (2002) Sequencing and transcriptional analysis of the chlorite dismutase gene of Dechloromonas agitata and its use as a metabolic probe. *Appl Environ Microbiol* **68**: 4820–4826.

Bender, K.S., Rice, M.R., Fugate, W.H., Coates, J.D., and Achenbach, L.A. (2004) Metabolic primers for detection of (Per)chlorate-reducing bacteria in the environment and phylogenetic analysis of cld gene sequences. *Appl Environ Microbiol* **70**: 5651–5658.

Bruce, R.A., Achenbach, L.A., and Coates, J.D. (1999) Reduction of (per)chlorate by a novel organism isolated from paper mill waste. *Environ Microbiol* **1**: 319–329.

Chaudhuri, S.K., O'Connor, S.M., Gustavson, R.L., Achenbach, L.A., and Coates, J.D. (2002) Environmental factors that control microbial perchlorate reduction. *Appl Environ Microbiol* **68**: 4425–4430.

Coates, J. and Achenbach, L. (2004) Microbial perchlorate reduction: rocket-fueled metabolism. *Nature Reviews Microbiology* **2**: 569–580.

Coates, J.D., Chakraborty, R., Lack, J.G., O'Connor, S.M., Cole, K.A., Bender, K.S., and Achenbach, L.A. (2001) Anaerobic benzene oxidation coupled to nitrate reduction in pure culture by two strains of Dechloromonas. *Nature* **411**: 1039–1043.

Coates, J.D., Michaelidou, U., Bruce, R.A., O'Connor, S.M., Crespi, J.N., and Achenbach, L.A. (1999) Ubiquity and diversity of dissimilatory (per)chlorate-reducing bacteria. *Appl Environ Microbiol* **65**: 5234–5241.

Domínguez, N.M., Hackett, K.T., and Dillard, J.P. (2011) XerCD-mediated site-specific recombination leads to loss of the 57-kilobase gonococcal genetic island. *J Bacteriol* **193**: 377–388.

Eddy, S. (1998) Profile hidden Markov models. *Bioinformatics (Oxford, England)* **14**: 755–763.

Fukuda, Y., Okamura, Y., Takeyama, H., and Matsunaga, T. (2006) Dynamic analysis of a genomic island in Magnetospirillum sp. strain AMB-1 reveals how magnetosome

synthesis developed. *FEBS letters* **580**: 801–812.
Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic biology* **52**: 696–704.
Ippen-Ihler, K.A. and Minkley, E.G. (1986) The conjugation system of F, the fertility factor of Escherichia coli. *Annu. Rev. Genet.* **20**: 593–624.
Juhas, M., der Meer, van, J.R., Gaillard, M., Harding, R.M., Hood, D.W., and Crook, D.W. (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS microbiology reviews* **33**: 376–393.
Kirk, A.B., Martinelango, P.K., Tian, K., Dutta, A., Smith, E.E., and Dasgupta, P.K. (2005) Perchlorate and iodide in dairy and breast milk. *Environ Sci Technol* **39**: 2011–2017.
Kounaves, S.P., Stroble, S.T., Anderson, R.M., Moore, Q., Catling, D.C., Douglas, S., et al. (2010) Discovery of natural perchlorate in the Antarctic Dry Valleys and its global implications. *Environ Sci Technol* **44**: 2360–2364.
Lawrence, J.E., Lamm, S.H., Pino, S., Richman, K., and Braverman, L.E. (2000) The effect of short-term low-dose perchlorate on various aspects of thyroid function. *Thyroid* **10**: 659–663.
Lima, T., Auchincloss, A., Coudert, E., Keller, G., Michoud, K., Rivoire, C., et al. (2009) HAMAP: a database of completely sequenced microbial proteome sets and manually curated microbial protein families in UniProtKB/Swiss-Prot. *Nucleic acids research* **37**: D471–8.
McMurdie, P.J., Behrens, S.F., Müller, J.A., Göke, J., Ritalahti, K.M., Wagner, R., et al. (2009) Localized plasticity in the streamlined genomes of vinyl chloride respiring Dehalococcoides. *PLoS genetics* **5**: e1000714.
McMurdie, P.J., Hug, L.A., Edwards, E.A., Holmes, S., and Spormann, A.M. (2011) Site-Specific Mobilization of Vinyl Chloride Respiration Islands by a Mechanism Common in Dehalococcoides. *BMC Genomics* **12**: 287.
Melnyk, R.A., Engelbrektson, A., Clark, I.C., Carlson, H.K., Byrne-Bailey, K., and Coates, J.D. (2011) Identification of a perchlorate reduction genomic island with novel regulatory and metabolic genes. *Appl Environ Microbiol* **77**: 7401–7404.
Sanchez, C.A., Crump, K.S., Krieger, R.I., Khandaker, N.R., and Gibbs, J.P. (2005) Perchlorate and nitrate in leafy vegetables of North America. *Environ Sci Technol* **39**: 9391–9397.
Semsey, S., Blaha, B., Köles, K., Orosz, L., and Papp, P.P. (2002) Site-specific integrative elements of rhizobiophage 16-3 can integrate into proline tRNA (CGG) genes in different bacterial genera. *J Bacteriol* **184**: 177–182.
Thrash, J.C., Ahmadi, S., Torok, T., and Coates, J.D. (2010) Magnetospirillum bellicus sp. nov., a novel dissimilatory perchlorate-reducing alphaproteobacterium isolated from a bioelectrical reactor. *Appl Environ Microbiol* **76**: 4730–4737.

van Ginkel, C.G., Rikken, G.B., Kroon, A.G.M., and Kengen, S.W.M. (1996) Purification and characterization of chlorite dismutase: a novel oxygen-generating enzyme. *Arch Microbiol* **166**: 321–326.

Tables and Figures

Table 1. *Characteristics of the four sequenced DPRB genomes and their PRIs*

Organism	***Dechloromonas aromatica* RCB**	***Azospira suillum* PS**	***Dechloromonas agitata* CKB**	***Magnetospirillum bellicus* VDY**
Genome Status	Finished	Draft (22 contigs)	Draft (120 contigs)	Draft (324 contigs)
Taxonomy (Class)	Betaproteobacteria	Betaproteobacteria	Betaproteobacteria	Alphaproteobacteria
Taxonomy (Order)	Rhodocyclales	Rhodocyclales	Rhodocyclales	Rhodospirillales
Taxonomy (Family)	Rhodocyclaceae	Rhodocyclaceae	Rhodocyclaceae	Rhodospirillaceae
Genome Size	4.5 Mb	3.8 Mb	4.0 Mb	8.4 Mb
Number of genes	4171	3454	3957	8040
Size of PRI 'core'	19.1 kb	19.5 kb	10.3 kb	17.6 kb
GC Content of Chromosome	59.2%	65.3%	62.9%	61.5%
GC Content of PRI 'core'	50.9%	51.2%	53.4%	63.2%
chlorite dismutase (*cld)*	Yes	Yes	Yes	Yes
perchlorate reductase (*pcrABCD)*	Yes	Yes	Yes	Yes (*pcrABD* duplicated)
Mo cofactor gene (*moaA)*	Yes	Yes	Yes	Yes
Two-component system	Yes	Yes	No	Yes (missing PAS domain protein)
Sigma-factor/anti-sigma factor	Yes	Yes	Yes	No
Unknown oxidoreductases	Yes	Yes	No	Yes (unrelated to RCB/PS genes)
Unknown c-type cytochromes	Yes	Yes	Yes	Yes
FNR/RpoB *pcrA* promoter	Yes	Yes	No	No

Figure 1. *Phylogenetic distribution of genomes of perchlorate reducers.*

This tree is the result of multilocus sequence analysis using a concatenated alignment of 9 conserved housekeeping proteins (AtpD, DnaA, DnaK, FtsZ, GyrB, RecA, RpoB, SecA, and Srp54). The parameters for this alignment were optimized using ProtTest and the tree was calculated with bootstrap values by phyml (Guindon and Gascuel, 2003; Abascal et al., 2005). The names of perchlorate reducers are identified with a gray background.

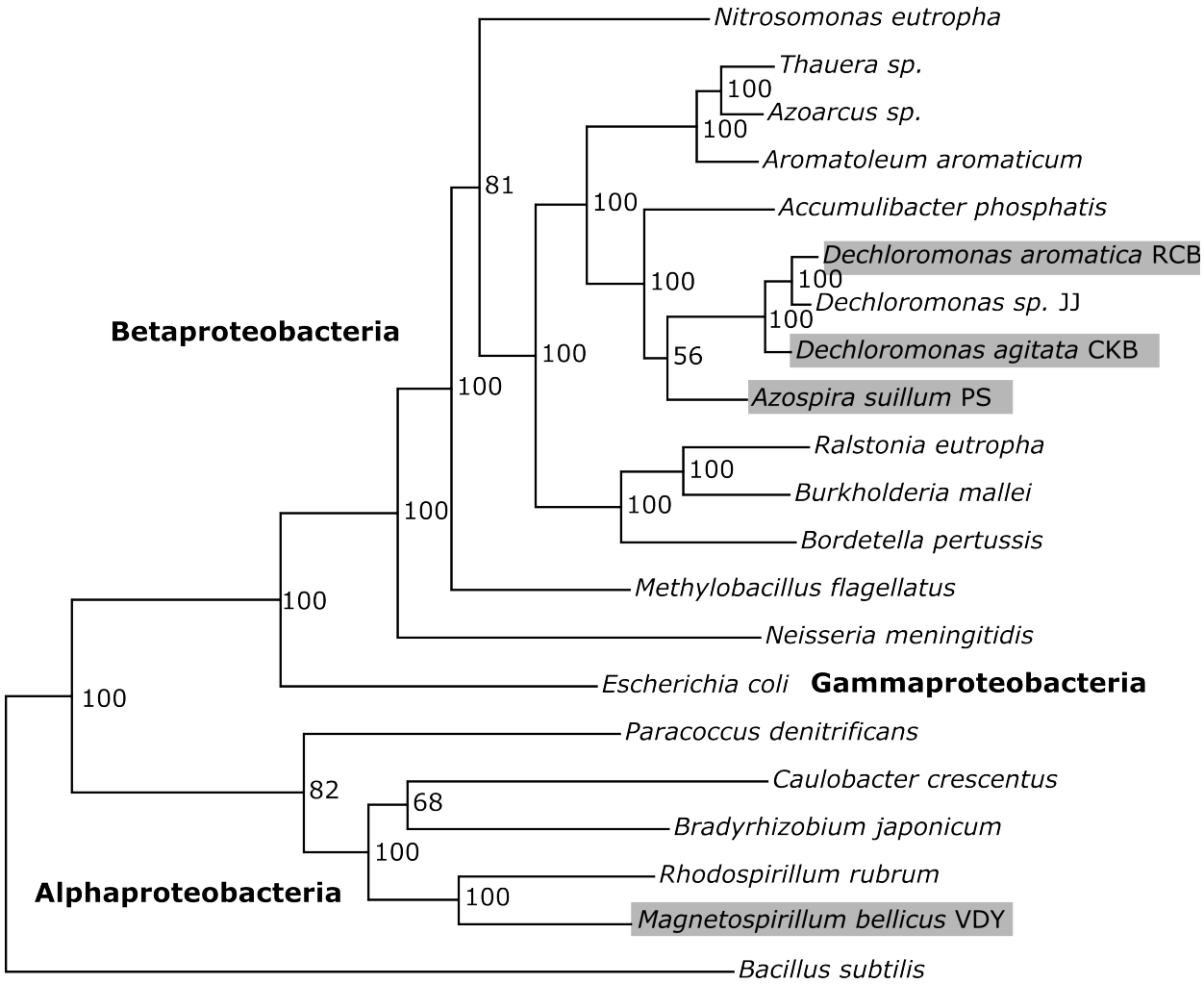
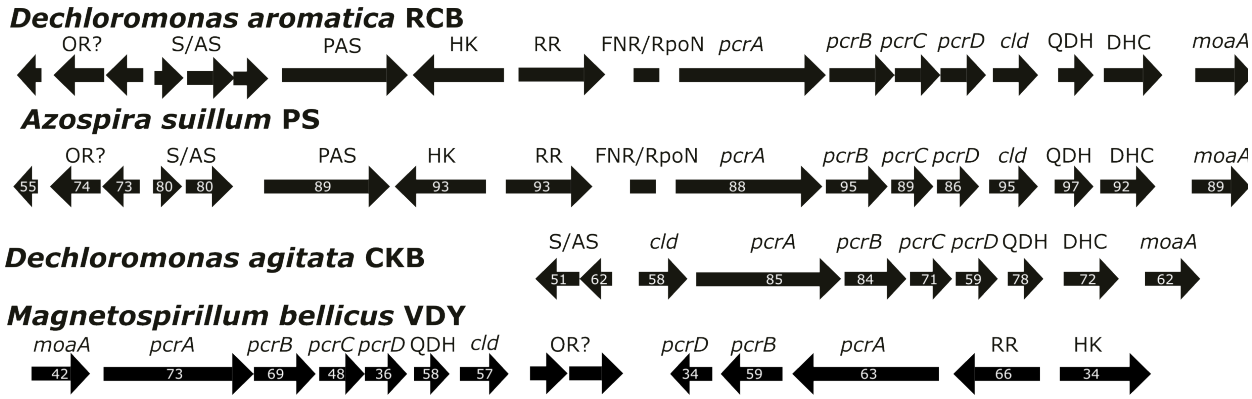


Figure 2. *Structure of the conserved 'core' of the perchlorate reduction genomic island (PRI).* The conserved genes from four examples of the PRI are shown here, using the conserved *pcrABCD* operon structure to generally align the four islands. The genes are labeled either with their given gene name, or an abbreviation/acronym for the predicted function of the gene product. The white numbers represent amino acid similarity relative to the homologous gene in the *D. aromatica* PRI. *pcrABCD* are components/accessory genes of perchlorate reductase, *cld* is chlorite dismutase, and *moaA* is a part of the molybdenum cofactor biosynthesis pathway. QDH denotes a gene predicted to encode a membrane-associated tetraheme *c*-type cytochrome with quinol dehydrogenase activity, while DHC denotes a diheme-*c*-type cytochrome. HK, RR, and PAS represent the histidine kinase, response regulator, and PAS domain sensor of a putative two-component system, while S/AS is a predicted sigma factor/anti-sigma factor system. OR? indicates genes annotated as various types of oxidoreductase components, but the substrate and function of their gene products within the context of perchlorate reduction is unknown. FNR/RpoN represents a conserved promoter for *pcrA* that contains consensus binding sites for FNR and RpoN.



Chapter 2:

Transposon and deletion mutagenesis of genes involved in perchlorate reduction in *Azospira suillum* PS

Abstract

Although much work has been published on the biochemistry of the key enzymes of bacterial perchlorate reduction, chlorite dismutase and perchlorate reductase, understanding of the molecular mechanisms of this metabolism has been somewhat hampered by the lack of a clear model system amenable to genetic manipulation. Using transposon mutagenesis and clean deletions, genes important for perchlorate reduction in *Azospira suillum* PS have been identified both inside and outside the previously described perchlorate reduction genomic island (PRI). Transposon mutagenesis identified 18 insertions in 11 genes that completely abrogate growth via reduction of perchlorate, but have no phenotype during denitrification. Of the mutants deficient in perchlorate reduction, 14 had insertions that were mapped to eight different genes within the PRI, highlighting its importance in this metabolism. To further explore the role of these genes, we also developed systems for constructing unmarked deletions and for complementing these deletions. Using these tools, every core gene in the PRI was systematically deleted; eight of the 17 genes conserved in the PRI are essential for perchlorate respiration, including three genes that comprise a unique histidine kinase system. Interestingly, the other nine genes in the PRI are not essential for perchlorate reduction and may thus have unknown functions during this metabolism. We present a model detailing our current understanding of perchlorate reduction that incorporates new concepts about this metabolism.

This chapter was originally published as an article in the journal *mBio* under the following reference: (Melnyk *et al.* (2014), **5**:ee0769-13).

Importance

Although perchlorate is generated naturally in the environment, groundwater contamination is largely a result of industrial activity. Bacteria capable of respiring perchlorate and remediating contaminated water have been isolated, but relatively little is known about the biochemistry and genetics of this process. Here we use two complementary approaches to identify genes involved in perchlorate reduction. Most of these genes are located on a genomic island, which is potentially capable of moving between organisms. Some of the genes identified are known to be directly involved in the metabolism of perchlorate, but other new genes likely regulate the metabolism in response to environmental signals. This work has uncovered new questions about the regulation, energetics, and evolution of perchlorate reduction, but also presents the tools to address them.

Introduction

Over billions of years, bacteria have evolved the ability to reduce dozens of different electron acceptors as part of energy metabolisms, from widely distributed molecules such as oxygen, nitrate, and sulfate to much rarer compounds like perchlorate. Naturally occurring perchlorate is thought to have an atmospheric origin but is deposited only in very small quantities (Rajagopalan et al., 2009); it accumulates mainly in hyperarid environments such as the Atacama Desert and the Antarctic Dry Valleys (Kounaves et al., 2010). Anthropogenic perchlorate, however, is a byproduct of various industries and can lead to localized contamination of groundwater due to its chemical stability and aqueous solubility (Parker, 2009). Despite the sporadic accumulation of perchlorate in the modern terrestrial environment, dissimilatory perchlorate-reducing bacteria (DPRB) are ubiquitous and can be isolated from many different environments (Coates et al., 1999).

Perchlorate reduction by microorganisms has been observed for decades but was initially thought to be a product of the nitrate reduction machinery common to many eubacteria and archaea (Coates and Achenbach, 2004). Further study of pure culture DPRB isolates revealed two enzymes specific to this metabolism: perchlorate reductase (PcrAB) and chlorite dismutase (Cld). PcrAB is a heterodimeric molybdoenzyme that belongs to the widespread and functionally diverse DMSO reductase family (Bender et al., 2005). PcrAB reduces perchlorate to chlorite, which is then dismutated into chloride and oxygen by the heme-containing Cld (van Ginkel et al., 1996; Kengen et al., 1999). The chlorite dismutase gene was initially identified in the chlorate reducer *Ideonella dechloratans* and subsequently in the DPRB *Dechloromonas agitata* CKB (Thorell et al., 2002; Bender et al., 2002). In the regions flanking the *cld* gene in two *Dechloromonas* species, the genes for the molybdopterin-containing perchlorate reductase were found (Bender et al., 2005). More recently, the entire genomes of several DPRB have been

sequenced, and all share a perchlorate reduction genomic island (PRI) that includes *pcrAB* and *cld* in addition to several conserved genes with unknown function (Melnyk et al., 2011). Similarly, a comparative study of chlorate reducers found that chlorate reduction genes are associated with composite transposons which contain conserved genes with unknown function (Clark et al., 2013). These comparative genomics studies provided evidence that (per)chlorate reduction genes are transferred horizontally, and several new genes were identified in the PRI that could be involved in regulation of the metabolism (Melnyk et al., 2011). Previous efforts to understand the regulation of perchlorate reduction by environmental conditions have used the enzymatic activity of whole cells or extracts, but little is known about the genetic or transcriptional basis for these observations (Chaudhuri et al., 2002; Xu et al., 2004).

While continued isolation of novel DPRB provides a diverse collection of bacteria capable of this metabolism (Thrash, Pollock, et al., 2010; Thrash, Ahmadi, et al., 2010; Nadaraja et al., 2012), understanding the basic tenets of perchlorate reduction is hindered by the lack of a robust model system amenable to genetic manipulation. To fill this role, we have selected the Betaproteobacterium *Azospira suillum* PS (formerly known as *Dechlorosoma suillum)* (Achenbach et al., 2001). PS is a facultative aerobe that is capable of respiring perchlorate and nitrate under anaerobic conditions. It grows rapidly and robustly on rich media and its genome has been fully sequenced and is publicly available (Melnyk et al., 2011; Byrne-Bailey and Coates, 2012). Additionally, its perchlorate-reducing physiology has been studied in the context of environmental regulation (Chaudhuri et al., 2002), chemotaxis (Sun et al., 2009), and redox cycling of electron shuttles (Van Trump and Coates, 2009; Clark et al., 2012).

In this study, we set out to broaden our knowledge about the molecular mechanisms and genetic factors involved in perchlorate reduction. We performed a random transposon mutagenesis of PS and identified 18 mutants unable to grow on perchlorate, most of which contained transposon insertions in the PRI. To further dissect this genomic island, we developed a method for creating markerless deletions and constructed in-frame knockout mutants of the 17 core genes in the PRI. We found that eight were essential for perchlorate reduction, while the remaining nine genes were either partially defective during growth on perchlorate or had no obvious phenotype. As expected, *pcrABCD* and *cld* were essential for perchlorate reduction, but so were genes for a response regulator (*pcrR*), histidine kinase sensor (*pcrS*), and a PAS domain-containing protein (*pcrP*), which form a putative regulatory system. Complementation of all eight genes using plasmid-encoded copies of the deleted gene resulted in partial or complete restoration of growth via perchlorate reduction. This study expands the list of genes essential for perchlorate reduction, but also supplies the tools for further analysis of the genetics and regulation of this metabolism.

Results

Transposon mutagenesis of PS and screening for mutants

Transposon mutagenesis and screening identifies the genetic factors involved in a given phenotype in an unbiased manner. We mutagenized PS using a plasmid-based MiniHimar transposon, which was delivered using the conjugation-competent *E. coli* strain WM3064 (Larsen et al., 2002; Bouhenni et al., 2005). Mutants were isolated as single, kanamycin-resistant colonies aerobically on rich ALP media to maximize the breadth of genes disrupted during mutagenesis. In total, the transposon library consisted of 2600 mutants.

To screen the mutants for defects associated with perchlorate reduction, we took advantage of the fact that PS will grow rapidly in the presence of both nitrate and perchlorate when transferred from aerobic conditions (Figure 1). Additionally, when both nitrate and perchlorate are present under anaerobic conditions, nitrate is used in preference to perchlorate by this organism, resulting in a diauxic growth curve (Chaudhuri et al., 2002). The mutant library was screened for growth under these conditions over a 30-hour period. Individual mutants that exhibited defects in either growth rate or final optical density were restreaked and selected for further study. To identify mutants with perchlorate-specific defects, each mutant was further characterized for the ability to grow on nitrate, perchlorate, nitrate plus perchlorate, and aerobically. Many mutants that were selected for additional characterization displayed wild-type growth characteristics, suggesting that the initial screen had a high incidence of false positives. Some mutants were pleiotropic, suffering defects of varying intensity in both growth rate and final optical density under multiple conditions. In order to focus on mutants directly involved in perchlorate reduction, we defined a subset of the mutants as 'perchlorate null' mutants. These mutants did not grow at all on perchlorate, but exhibited normal growth on nitrate alone, as well as normal aerobic growth on both solid and liquid ALP media. The full phenotypic characterizations of the transposon mutants of interest are detailed in Table 1. In total, 18 perchlorate null mutants and 18 pleiotropic mutants were isolated (Table 1).

Transposon mutagenesis highlights the importance of genes in the PRI

The genomic locations of transposon insertions were identified using an arbitrary PCR-based method (Table 1). Many pleiotropic mutants were identified in genes predicted to be important for anaerobic respiration, including molybdopterin biosynthesis (*moeA* and *mogA*), molybdenum uptake (*modC*), nitrate reduction/denitrification (*napA*, *norD*, and Dsui_1177), anaerobic heme biosynthesis (*hemN*), and ubiquinone biosynthesis (*ubiX*). Additionally, three insertions were identified in the *rnf* operon that resulted in phenotypes that were comparable to wild-

type when grown on nitrate, but grew very slowly on perchlorate. Similarly, an insertion in a gene from the *relA*/*spoT* family (Dsui_2816) had a very slow growth rate on perchlorate.

While the pleiotropic mutants had insertions in genes located throughout the chromosome, most (14 out of 18) of the perchlorate null mutant insertions were localized to eight genes in the PRI (Melnyk et al., 2011). Insertions were found in previously identified structural genes (*pcrA, pcrB, pcrD,* and *cld*), but also in several genes of unknown function. Three genes in the PRI identified in this screen encode a putative histidine kinase signal transduction system (HKS). We have named these genes *pcrR*, *pcrS,* and *pcrP* to reflect their linkage with *pcrA* and the fact that homologs of some or all of these genes can be identified in the PRI of other perchlorate reducers, some of which are distantly related to PS (Melnyk et al., 2011). The gene names *pcrR*, *pcrS,* and *pcrP* are intended to be suggestive of the putative roles of each gene product in the HKS: the response **R**egulator, the kinase **S**ensor, and the **P**AS domain-containing protein. PcrP and PcrS are predicted to be integral membrane proteins with large regions in both the cytoplasm and periplasm. The cytoplasmic region of PcrS contains domains seen in many bacterial histidine kinases: a HAMP linker domain (PF00672), a phosphoacceptor domain (PF00512), and an ATPase domain (PF02518) (Galperin et al., 2001). PcrP contains two PAS domains from different subfamilies in its cytoplasmic region(PF08448 and PF13188) (Möglich et al., 2009). In contrast to the cytoplasmic regions which are composed of well-characterized and common domains, the periplasmic regions of PcrP and PcrS are not homologous to any domains in the COG or Pfam databases. These regions are fairly large, consisting of approximately 250 and 270 amino acids in PcrS and PcrP, respectively. BLAST-P searches of the PcrP domain against finished genomes identified several similar sequences, while a search using the PcrS domain identified only the PcrS homolog from another DPRB, *Dechloromonas aromatica* RCB. PcrR contains three domains commonly found in response regulators: the receiver domain (PF00072), a σ_{54} interaction domain (PF00158), and a helix-turn-helix DNA-binding domain (PF02954) (Galperin, 2006). Another insertion in a perchlorate null mutant mapped to the cupin domain protein gene (Dsui_0153) in the PRI. Outside of the PRI, three genes also contained insertions that led to the perchlorate null phenotype. Two separate mutants were isolated with insertions in Dsui_0704, which encodes the alternative σ_{54}-type sigma factor RpoN. The other two genes (Dsui_0128 and Dsui_1441) are of unknown function.

Two diauxic phenotypes associated with 'perchlorate null' mutants

Perchlorate null mutants were defined as having no growth on media containing perchlorate, while being indistinguishable from wild-type PS on media with nitrate. When the 18 perchlorate null mutants were grown on diauxic media containing nitrate

and perchlorate, two different phenotypes were observed. Most (15 out of 18) of the perchlorate null mutants had similar growth curves on diauxic media and nitrate-only media. The remaining three perchlorate null mutants had significantly impaired growth on diauxic media relative to media containing only nitrate. Interestingly, all three of these mutants had transposon insertions in *cld*, the gene encoding chlorite dismutase (Figure 2). The other 15 perchlorate null mutants had insertions in ten different genes including *pcrA* (Figure 2). To highlight the different diauxic growth phenotypes, *pcrA*::Himar and *cld*::Himar transposon mutants were compared to a biotin biosynthetic mutant (*bioA*::Himar) which had wild-type growth characteristics on kanamycin-amended ALP media, presumably due to the presence of biotin in the vitamin supplement (Figure 3A-C). Neither the *pcrA*::Himar or *cld*::Himar mutants grew at all on perchlorate (Figure 3B), but the *cld*::Himar mutant grew to a much lower optical density than the *pcrA*::Himar mutant on nitrate and perchlorate (Figure 3C).

Developing a genetic system in PS and systematic deletions of genes in the PRI

To confirm that genes identified in the transposon screen were essential for perchlorate reduction, we developed a system for making markerless chromosomal deletions in PS. Several different suicide vectors did not undergo chromosomal integration when delivered via conjugation, but a suicide vector derived from pNPTS138 (Dickon Alley, unpublished data) was able to integrate into the chromosome. However, conjugation proved to be an unpredictable method of vector delivery, as conjugations would frequently fail and no chromosomal insertions could be recovered. Reliable chromosomal insertions of vectors derived from pNPTS138 were achieved using electroporation, although efficiency of integration was low (five to ten insertions per μg pDNA transformed). This is consistent with the previous observation that *Azospira* spp. are resistant to transformation (Reinhold-Hurek and Hurek, 2006).

We used pNPTS138 to create deletions of the 17 genes in the conserved core of the PRI. Because many genes in the PRI are in operons (e.g. *pcrABCD*), we designed the deletion constructs to create in-frame deletions of each gene that did not disturb the ribosomal binding site or promoters of nearby genes. Once each deletion was made, we characterized its ability to grow on nitrate, perchlorate, and nitrate plus perchlorate. Of the eight genes in the PRI that were defined as being perchlorate null mutants during the transposon screen, seven (*pcrABD*, *cld*, and *pcrPSR*) recapitulated the perchlorate null phenotype. Additionally, the deletion of *pcrC* resulted in a perchlorate null phenotype, although no insertions in *pcrC* were isolated during the transposon screen. When Dsui_0153 was deleted, the resulting strain exhibited no differences from wild-type under all conditions tested, suggesting that the perchlorate null phenotype of Dsui_0153::Himar was the result of a polar effect, possibly on the adjacent *pcrP* gene.

In addition to Dsui_0153, six other genes in the PRI also had no apparent phenotype upon deletion. Five of these genes (Dsui_0154-Dsui_0158) are situated at one end of the PRI and comprise two oppositely oriented operons. The five genes contain a predicted sigma factor/anti-sigma factor pair, genes homologous to *yedYZ* from *E. coli* (Loschi et al., 2004), and a gene encoding a hypothetical protein. The sixth non-essential gene (Dsui_0141) is located at the opposite end of the PRI and is a homolog of *moaA*, a gene involved in molybdopterin biosynthesis (Schwarz et al., 2009). Two additional genes in the PRI, Dsui_0143 and Dsui_0144, were not essential for perchlorate reduction, but both suffered from slight yet significant growth rate defects on perchlorate. Dsui_0144 is predicted to encode a tetraheme *c*-type cytochrome homologous to proteins such as NapC, NirT, and CymA, all of which function as quinol dehydrogenases in electron transport chains involved in the reduction of various electron acceptors (Simon and Kern, 2008). Dsui_0143 is predicted to be a gene encoding a protein homologous to EbdC, the gamma subunit of ethylbenzene dehydrogenase of *Aromatoleum aromaticum*, which contains a heme *b* and forms a heterotrimer with the alpha and beta subunits (Kloer et al., 2006). However, the Dsui_0143 protein may contain two additional heme *c* molecules as it contains two putative heme *c* binding motifs (one CXXCH and an atypical CWXXCH). These two motifs are conserved in the homologous NirB protein which is in a cluster of genes involved in nitrite reduction in *Pseudomonas stutzeri* (Jüngst et al., 1991). The growth phenotypes of all 17 single deletions are summarized in Table 2.

Complementation of perchlorate reduction mutants

To show that the eight genes from the PRI were independently essential for perchlorate reduction and the observed phenotypes were not the result of polar effects, we developed a system for complementation analysis. The broad-host range plasmid pBBR1MCS2 was tested for stable replication in PS and was used as the backbone for all complementation vectors. The complementation vectors were constructed in two groups. The first group contained the genes *pcrPSR*, which are assumed to be transcribed independently based on their alternating orientation (Figure 2). When constructing these vectors, we included the ~150 bp upstream of the start codon in order to retain the native promoter and regulation. The second group contains five essential genes (*pcrABCD* and *cld*) located in the same operon. To construct these vectors we first made a plasmid with the *pcrA* promoter (pRAM62) and then incorporated each of the five genes independently into pRAM62 with their endogenous ribosomal binding site. Each of the eight perchlorate null deletion mutants received both the complementation vector and the empty pBBR1MCS2 vector. Under anaerobic conditions with both nitrate and perchlorate, all eight complemented strains were able to grow on perchlorate, while the strains with the empty vector recapitulated the defective growth

curves seen in the deletion mutants (Figures 4-6). The complemented strains had variable lag times when switching from nitrate to perchlorate and different growth rates on perchlorate. The reasons for this are unknown, but could be due to differences in expression between the complemented mutants and wild-type PS.

Phenotypic characterization of the Δ*cldΔpcrA *double mutant

To further explore the discrepancy between the diauxic phenotype of Δ*cld* and Δ*pcrA*, we constructed a Δ*cld*Δ*pcrA* double mutant by electroporating the *cld* deletion vector into the previously constructed Δ*pcrA* strain. As previously described, the Δ*cld* strain exhibited normal growth on nitrate alone, but was impaired for growth on nitrate when perchlorate was also present (Figure 7A). We hypothesized that this was due to baseline expression of *pcrA* and accumulation of chlorite in the absence of Cld activity. As expected, the Δ*cld*Δ*pcrA* strain did not grow on perchlorate, but was able to grow normally on nitrate (Figure 7A, 7B). However, in the presence of nitrate and perchlorate, the Δ*cld*Δ*pcrA* mutant recovered the ability to grow on nitrate like the Δ*pcrA* single deletion (Figure 7C).

Discussion

Previous work identified the *cld* gene by sequencing a fragment of a purified enzyme with chlorite dismutase activity (Stenklo et al., 2001). This sequence was used to amplify and sequence *cld* from *Dechloromonas agitata* CKB and subsequently used to identify neighboring perchlorate reductases (Bender et al., 2002; 2005). *pcrA* was directly mutagenized and shown to be essential in *Dechloromonas aromatica* RCB, however, this was done with an antibiotic cassette insertion and may have resulted in polar effects on nearby genes (Bender et al., 2005). A forward genetics approach confirmed that *cld* and *pcrA* are essential for perchlorate reduction in *Azospira suillum* PS, and also identified the importance of *pcrB, pcrD,* a new HKS, and *rpoN*. Markerless in-frame deletion were used to confirm the insertions in the PRI and also led to the identification *pcrC* as an essential gene for perchlorate reduction. These complementary genetic approaches advance our understanding of the genetic factors involved in perchlorate reduction and establish *Azospira suillum* PS as a model perchlorate reducer.

The majority of the genes identified as essential for perchlorate reduction are located in the perchlorate reduction genomic island (PRI), which contains 17 genes in PS that are conserved in *Dechloromonas aromatica* RCB but not in other closely related organisms incapable of perchlorate reduction (Melnyk et al., 2011). Our transposon screen identified *cld, pcrA*, *pcrB,* and *pcrD* as essential for perchlorate reduction, which we confirmed with deletion mutagenesis. Although it was not identified in the transposon screen, *pcrC* was also found to be essential for perchlorate reduction. *pcrB* encodes the beta subunit of the perchlorate reductase and the initial biochemical

characterization of perchlorate reductase from *Azospira* sp. GR-1 identified a stable enzyme complex in a $\alpha_3\beta_3$ conformation (Kengen et al., 1999). PcrB is likely required for the reductase complex stability and for delivering reducing equivalents from the inner membrane electron transport chain to the active site of PcrA. PcrC is a multiheme cytochrome *c* that was previously thought to be part of the perchlorate reduction pathway in all DPRB (Bender et al., 2005), but a recently isolated and described organism, *Arcobacter* sp. CAB, reduces perchlorate and does not have a gene homologous to *pcrC* in its PRI or elsewhere in its genome (Carlström et al., 2013). *pcrD* encodes a putative chaperone that is homologous to both TorD and NarJ of *E. coli* and has been predicted to be the chaperone for PcrA (Blasco et al., 1998; Jack et al., 2004; Bender et al., 2005). Many periplasmic molybdoenzymes of the DMSO reductase family have dedicated chaperones which bind to the signal peptide and prevent transport of the apoprotein via the Tat (**T**win-**a**rginine **t**ranslocation pathway) until addition of the appropriate cofactor (Jack et al., 2004; Maillard et al., 2007). The perchlorate null phenotype of the *ΔpcrD* strain suggests that PcrD may similarly be required for the attachment of molybdopterin to PcrA and the proper translocation of the holoenzyme.

pcrP, *pcrR,* and *pcrS* were identified in both the transposon screen and deletion mutagenesis as being essential for perchlorate reduction. Together, their gene products comprise a histidine kinase signal transduction system that is conserved in other DPRB (Melnyk et al., 2011). The sensory inputs and physiological outputs of this system are currently unknown, but the DNA-binding domain of PcrR suggests that it likely binds to promoters to modulate transcription. PcrR also contains a σ_{54} interaction domain, which can recruit σ_{54}-type sigma factors to promoters where the regulator is bound and hydrolyze ATP to initiate transcription (Bush and Dixon, 2012). PS contains exactly one σ_{54}-type sigma factor encoded by the *rpoN* gene (Dsui_0704), which was one of only three genes outside of the PRI that resulted in a perchlorate null phenotype when mutated with a transposon insertion. Additionally, there is a conserved promoter region upstream of *pcrA* that contains a RpoN-binding motif (Melnyk et al., 2011). Based on these observations, we hypothesize that at least one of the outputs of this system is to upregulate transcription of the *pcrABCD* operon in an RpoN-dependent manner. The presence of two separate sensor proteins that are both essential (PcrP and PcrS) hints at a complex regulatory mechanism.

The periplasmic domains of PcrP and PcrS are unique and the signals that modulate this system are unknown. Perchlorate, nitrate, or intermediates of both reductive pathways could be sensed directly, as is the case for the *E. coli* NarX sensor kinase, which autophosphorylates in response to binding nitrate at the dimerization interface (Lee et al., 1999). An emerging paradigm for histidine kinase "two-component" signal transduction is the presence of a third auxiliary protein which can adjust the autophosphorylation of the histidine kinase and/or phosphotransfer to the response

regulator (Buelow and Raivio, 2010). PcrP could be such an auxiliary protein and it could sense signals that originate either in the periplasm or in the cytoplasm, perhaps via its two PAS domains. Interestingly, *pcrPSR* do not appear in the genome of *Dechloromonas agitata* CKB, a DPRB that does not respire nitrate unlike PS, suggesting that the PRI HKS could play a role in nitrate-dependent repression of perchlorate reduction (Bruce et al., 1999; Chaudhuri et al., 2002; Melnyk et al., 2011). We have presented a schematic for the basic functions of the PRI HKS, highlighting the domain structure of PcrP and PcrS in the inner membrane, as well as the putative function of PcrR in initiating transcription in an RpoN-dependent manner (Figure 8). Although we currently know little about how the system functions, molecular analysis of the signals and domains involved will be crucial in understanding the evolution of the PRI.

Nine genes in the PRI were not required for perchlorate reduction (Dsui_0141, Dsui_0143-Dsui_0144, and Dsui_0153-Dsui_0158). All of these genes are conserved in at least one other DPRB, thus it was unexpected that so many (9 out of 17) had no obvious role in perchlorate reduction (Melnyk et al., 2011). However, a similar observation was made in a genetic study of the magnetosome genomic island, where many of the conserved genes were not essential for magnetosome formation (Murat et al., 2010). Possible reasons proposed for this phenomenon were genetic redundancy, as well as the inherent limitations of laboratory-imposed growth conditions (Murat et al., 2010). It is reasonable to expect that both of these caveats are relevant for genetic analysis of the PRI.

The perchlorate reductase receives electrons from the electron transport chain via an unknown pathway. PS has a predicted quinol dehydrogenase-type cytochrome *c* (Dsui_0144) in its PRI that was previously proposed to be a conduit from the quinone pool to the reductase (Bender et al., 2005), but there was only a slight impairment of growth on perchlorate when the gene was deleted in PS. A similar phenotype was seen with the deletion of the predicted cytochrome *b*/*c* Dsui_0143. Because none of these genes are absolutely essential for perchlorate reduction, we propose that the electron transport chain to PcrA is redundant. This sort of multiplicity is seen on a larger scale in organisms such as *Geobacter* spp., which utilize dozens of different cytochrome *c* proteins to transport electrons across the cell envelope to multiple electron acceptors (Shi et al., 2007). The perchlorate electron transport chain may not be quite as complex, but additional genetic and biochemical evidence is needed to resolve this pathway in PS. The limitations of our experimental growth conditions may also occlude the phenotypes of certain genes in the PRI. For example, Dsui_0154 and Dsui_0155 are predicted to encode an extracytoplasmic function sigma factor/anti-sigma factor system homologous to SigF and NrsF of *Caulobacter crescentus*, which has been shown to play a role in activating a response to reactive oxygen species and heavy metal stress (Alvarez-Martinez et al., 2006; Kohler et al., 2012). The phenotype of these knockouts in PS may

only be apparent under the appropriate oxidative stress conditions which are not met by our standard acceptor-limited conditions in rich media.

One possible source of oxidative stress associated with perchlorate reduction could be an intermediate of the metabolism. Both the transposon screen and deletion mutagenesis revealed that the *cld* mutant had a different phenotype from the rest of the perchlorate null mutants when grown on nitrate and perchlorate together (Figure 3). Interestingly, deletion of *pcrA* in the *Δcld* background relieved the growth inhibition (Figure 7). In the context of perchlorate reduction, chlorite dismutase liberates molecular oxygen from chlorite and sustains the microaerobic respiration that is thought to accompany perchlorate reduction (Coates and Achenbach, 2004). However, Cld can also function to detoxify chlorite, which has been proposed as a role for Cld in organisms that can not reduce perchlorate or chlorate (Mlynek et al., 2011). Chlorite is a powerful oxidant that can result in cell death even at parts per million concentrations (Kwolek-Mirek et al., 2011). Previous experiments with PS indicate that perchlorate is not reduced with nitrate in large quantities, but if perchlorate reductase is expressed at all under nitrate-reducing conditions in the absence of Cld, the micromolar amounts of chlorite that accumulate may result in the inhibition of cell growth. A study performed with cell extracts of the closely related DPRB *Azospira* sp. KJ grown on nitrate showed some perchlorate reductase activity, suggesting that there could be baseline expression of PcrA even under nitrate-reducing conditions (Xu et al., 2004).

As previously proposed, perchlorate reduction is dependent on genes in the horizontally transferred PRI, but also on genes elsewhere in the genome (Melnyk et al., 2011). Our transposon screen identified several operons and pathways that are involved in anaerobic respiration. Insertions in *hemN* (Dsui_1647) and *ubiX* (Dsui_0993) resulted in pleiotropic mutants that were defective in growth on both nitrate and perchlorate. HemN is an oxygen-independent coproporphyrinogen III oxidase and UbiX is a decarboxylase involved in the biosynthesis of quinones (Troup et al., 1995; Gulmezian et al., 2007). Presumably, these mutants are deficient in anaerobic heme and quinone biosynthesis which leads to a general growth defect under anaerobic conditions. Perchlorate reductase requires a molybdopterin-bound molybdenum to function and several genes involved in molybdopterin biosynthesis were identified in the transposon screen (Table 1). These mutants were also wholly or partially defective during denitrification, which is likely due to the nitrate reductase utilizing the same cofactor. Additionally, the transposon screen allowed the identification of several genes putatively involved in denitrification, which has not been studied directly in PS. These genes include *napA, norD*, and Dsui_1177. The gene product of Dsui_1177 is a heme-containing protein which has been shown to be upregulated under denitrifying conditions (Clark et al., 2012). Interestingly, the *napA* and Dsui_1177 mutants also

demonstrate some defects on perchlorate reduction. The reason for this is currently unknown.

Many other genes were identified in the transposon screen as being partially defective in perchlorate reduction and not denitrification, but their specific function is difficult to predict. These include the *rnf* operon, which was hit by transposons several times in the mutagenesis and screen. The *rnf* operon encodes a membrane bound redox-active complex which couples the proton (or sodium) gradient of the inner membrane to the cellular redox pools of NAD^+/NADH and ferredoxin (Biegel et al., 2011). The activity of the Rnf complex is reversible and has been shown to be important in an array of different bacterial metabolisms, although its role in perchlorate reduction is currently unknown. Other transposon insertions are in genes that are not directly linked to cellular energetics, or are in genes that encode hypothetical proteins. Exploring the roles of these genes will be essential to gaining a holistic understanding of perchlorate reduction in the context of global cellular metabolism.

Conclusion

In this chapter, we set out to address a basic question about the molecular biology of perchlorate reduction: which genes are essential for this metabolism? Transposon mutagenesis and a subsequent screen on media containing nitrate and perchlorate identified many mutants with insertions in the PRI, accentuating its importance. To further characterize this region, we developed a system for making markerless in-frame deletions and systematically knocked out each gene in the PRI. The results of these complementary genetic approaches have increased our understanding of the factors involved in perchlorate reduction (Figure 8), but have raised many new questions. The histidine kinase system in the PRI is essential for perchlorate reduction, although we know little about how it is activated or which genes it regulates. Specifically, utilization of nitrate over perchlorate is still not understood. The components that deliver electrons to PcrA are also unknown, as is the role played by the non-essential genes in the PRI. Finally, the evolution of the PRI is still obscure; how does the island integrate with the host's metabolism after it is acquired? Genetic analysis using the deletion and complementation tools developed here will be critical to answering these questions about the evolution and function of the PRI in *Azospira suilllum* PS and thus other DPRB.

Materials and Methods

Bacterial strains and plasmids

Azospira suillum PS (ATCC BAA-33/DSMZ 13638) was revived from a lab freezer stock and used as the wild-type strain for all genetic manipulations (Table 3). Various *E. coli* strains were also used for cloning and conjugation purposes (Table 3). Plasmids constructed or used in this study are described in Table 4. Prior to all growth curves and genetic manipulations, all strains were streaked out from master freezer stocks to get single colonies.

Culture conditions and growth media

E. coli strains were grown in LB media. Kanamycin (Kan, 50 μg/mL) was used for selection and diaminopimelic acid (DAP, 0.3 mM) was used as a supplement to cultivate the auxotrophic strain WM3064. For routine culturing as well as growth assays, wild-type and mutant strains of PS were grown in ALP media. One liter of ALP media is composed of 0.49 g monobasic sodium phosphate dihydrate, 0.97 g dibasic anhydrous sodium phosphate, 0.1 g potassium chloride, 0.25 g ammonium chloride, 0.82 g sodium acetate, 2.0 g yeast extract, 7.6 g of a 60% w/w sodium lactate solution, 1.10 g sodium pyruvate and 10 mL of both vitamin mix and mineral mix as previously described (Bruce et al., 1999). To make solid ALP media for plates, 15 g/L agar was added. Kanamycin was used for selection of PS mutants at a concentration of 50 μg/mL. For anaerobic growth of PS, ALP media was supplemented with either 5 mM sodium nitrate, 2.5 mM sodium perchlorate, or 5 mM of both sodium nitrate and sodium perchlorate. No growth was observed when no electron acceptor was added, confirming that PS is unable to grow via fermentation of the carbon sources present in ALP media. All strains of *E. coli* and PS were cultivated at 37°C. Anaerobic growth curves of PS for the transposon screen and characterization of strains was performed using a Spectramax 340PC384 plate reader in an anaerobic chamber (Coy Laboratory Products, Grass Lake, MI).

Transposon mutagenesis via conjugation

The *mariner* transposon suicide vector pMiniHimar RB1 was used to mutagenize PS (Bouhenni et al., 2005). WM3064 containing pMiniHimar RB1 (Arkin Lab, UC Berkeley) was grown overnight culture in 5 mL LB+Kan+DAP. This culture was centrifuged and washed twice in ALP media to remove excess kanamycin. In parallel, a 5 mL culture of PS was grown from a single colony overnight (OD_{600} ~ 1.5) and centrifuged. The WM3064 and PS cell pellets were resuspended in 50 μL ALP media and spotted onto an ALP plate supplemented with DAP. The plate was dried by an open flame for 15 minutes on the benchtop to ensure that the agar surface was dry and the mating reaction did not spread. The plate was then incubated at 37°C for 6 hrs, after which the entire spot was scraped off the plate and resuspended in 1 mL ALP media. The volume was divided and plated on 6 ALP-Kan plates and incubated at 37°C. After

36-48 hours, ~300-500 whitish-pink opaque colonies formed on the 6 plates in total. The colonies were then transferred into 96-well plates containing 300 μL of ALP-Kan media supplemented with 7.5% glycerol as a cryoprotectant. The plates were then grown overnight aerobically on a platform shaker at 37°C prior to screening and long-term storage at -80°C.

Screening the transposon mutant library

The transposon library was screened prior to long-term storage. From each glycerol freezer stock plate, 10 μL was transferred into 300 μL of ALP-Kan media supplemented with 5 mM nitrate and perchlorate on the benchtop. Plates were then transferred into the anaerobic chamber, where they were allowed to equilibrate with the anaerobic atmosphere for 1 hour. An initial optical density reading was taken and the plates were placed in a GasPak Box (BD) with a palladium catalyst for maintenance of anaerobic conditions. The box was removed from the glove bag and incubated at 37°C. Every 6-8 hours for 30 hours, the box was moved back into the glove bag and the optical density of the plates was measured. The growth curve data was plotted and mutants that had a defect in either growth rate or maximum optical density were identified. The 'mutants of interest' were then streak-purified to isolate single colonies on ALP-Kan media. The single colonies were grown up on ALP media to make long-term freezer stocks.

Characterization of transposon mutants defective on perchlorate reduction

Mutants of interest identified in the initial screen were subjected to further phenotypic characterization. Each mutant was streaked out on ALP-Kan plates from freezer stocks and then transferred into an overnight culture of ALP-Kan liquid media. The next morning, the optical density of all strains was determined, and each culture was diluted back to an OD_{600} of 1.0-1.2. A 50 μL aliquot was used to inoculate 1.45 mL of ALP-Kan media supplemented with either nitrate (5 mM), perchlorate (2.5 mM), or nitrate and perchlorate (both 5 mM). The concentration of perchlorate in the perchlorate-only condition was lowered to avoid unpredictably long lag periods seen with high perchlorate concentrations (unpublished data). Each media condition was distributed to 4 wells in a 96-well plate such that 8 genotypes could be assayed simultaneously using the Spectramax plate reader in the anaerobic chamber. The characterization growth curves were allowed to run for 48 hours.

Mapping the location of transposon mutants

Transposon mutants of interest were mapped using a variation of arbitrarily primed PCR methods previously described (Jacobs et al., 2003; Das et al., 2005). Briefly, transposon mutants of interest were streaked onto ALP-Kan plates and single colonies

were used as templates for colony PCR. The first round of PCR used a primer that annealed to the end of the *mariner* transposon (HIMAR_EXT, Table 5) and one of three arbitrary primers (PS_ARB4, PS_ARB5, PS_ARB6) which contain a conserved 20-mer at the 5′ end, followed by 10 random nucleotides and one of three most common pentamers in the *A. suillum* PS genome. This first round of PCR used the GoTaq Green Master Mix (Promega) and the following thermocycling parameters: 10 minutes at 95°C; 35 cycles of 30 seconds at 95°C, 30 seconds at 38°C, 2 minutes at 72°C; and 10 minutes at 72°C. The reaction mixture was treated with ExoSAP-IT (Affymetrix) to remove the primers, and the product was used as the template for a second round of PCR using a transposon primer closer to the genomic insertion site (HIMAR_INT) and a primer consisting of the conserved 20-mer at the 5′ end of the arbitrary primers (PS_ARB2). The second round of PCR used the following thermocycling parameters: 5 minutes at 95°C; 35 cycles of 30 seconds at 95°C, 30 seconds at 55°C, 2 minutes at 72°C; and 10 minutes at 72°C. The second round PCR product was purified using the QIAquick PCR Purification Kit (Qiagen) and used as the template for a sequencing reaction with the HIMAR_INT primer (UC Berkeley DNA Sequencing Facility). Sequencing reads were manually curated and used as queries for searching the *A. suillum* genome with BLAST on the NCBI server (Altschul et al., 1990). Sequences were downloaded and the exact insertion site of the transposon in each mutant was determined.

Constructing suicide vectors and complementation vectors

Genomic DNA from *A. suillum* was isolated using Trizol according to the manufacturer's directions (Invitrogen) and all PCRs were performed using Phusion DNA Polymerase (Fisher Scientific). Specific primer pairs were designed using the NCBI Primer-BLAST server to amplify products of 800-900 bp flanking the gene to be deleted (Ye et al., 2012). The internal primers closest to the gene were designed to end between codons within the gene so that the resulting deletion would be in-frame and have minimal polar effects on nearby genes. For some genes, the internal primers had a 21 bp 'linker' at the 5′ end that was used to assemble the flanking regions into a 1.6-1.8 kb insert using PCR. For the remaining genes, the internal primers contained restriction sites that allowed assembly of the suicide vector with a three-way ligation. Both methods used resulted in an in-frame deletion allele of the gene of interest, thereby minimizing polar effects on nearby genes. Sequences and brief descriptions of all primers are listed in Table 5.

We used the vector pNPTS138 (Dickon Alley via Kathleen Ryan) as the suicide vector backbone for all of our deletion constructs after testing several plasmids with no success (pAK31, pK19mobsacB, and pSMV3). The broad host range plasmid pBBR1MCS2 was used as the backbone for the complementation vector (Kovach et al., 1995). Deletion vectors were constructed by amplification of two regions flanking the

gene of interest and either performing an assembly PCR to fuse the regions together or digesting the PCR product directly. In both cases, the cut vector and digested PCR product(s) were digested with the appropriate restriction enzymes, analyzed by electrophoresis on an agarose gel and extracted using the QIAEX II Gel Extraction Kit (Qiagen). The digested products were then ligated into the vector using T4 ligase (NEB). The broad host range plasmid pBBR1MCS2 was used as the backbone for the complementation vectors (Kovach et al., 1995). For three genes (*pcrP*, *pcrS*, and *pcrR*), pBBR1MCS2 and the amplified gene with its promoter were digested, purified, and ligated as described above. For *pcrABCD* and *cld* complementation vectors, the *pcrA* promoter was first cloned into pBBR1MCS2 to make pRAM62. The amplified genes were then cloned downstream of this promoter as described above to generate these complementation vectors. All ligation reactions were transformed into chemically competent *E. coli* TOP10 cells and plated onto selective medium. Individual colonies were grown overnight in liquid broth and plasmid DNA was purified. Plasmids were checked for presence of the correct insert by digestion with appropriate restriction enzymes, followed by sequencing of the insert (UC Berkeley DNA Sequencing Facility).

Electroporation of suicide vectors into PS and screening of deletions

Electrocompetent *A. suillum* PS cells were prepared using a method previously described for *Zymomonas mobilis* (Lam et al., 1993). 100 μL aliquots of electrocompetent cells were thawed on ice and mixed with 1 μg of plasmid DNA and 0.8 μL of TypeOne Restriction Inhibitor (Epicentre Biotechnologies). The mixture was transferred into a cuvette on ice and then electroporated at 1750 V, 400 Ω, and 25 μF. Immediately after electroporation, 300 μL of ALP media was added to the cuvette, and the entire aliquot was transferred to a sterile 1.7-mL tube for 6 hours of recovery in a 37°C shaking incubator. After the recovery, the aliquot was plated on 3 ALP-Kan plates and incubated at 37°C for 36 hours. Colonies that formed within 36 hours were picked into 500 μL of ALP-Kan liquid media and grown for 12 hours to confirm integration of the suicide vector into the chromosome. An aliquot of 50 μL was then transferred into 5 mL of ALP media and grown overnight. Dilutions of the overnight culture were plated on ALP plates with 6% sucrose (ALP-Suc) to select against the *sacB* allele present on the pNPTS138 backbone and incubated at 37°C for 36 hours. Colonies that appeared were patched onto ALP-Kan and ALP-Suc and incubated at 37°C overnight. Colonies that were sensitive to kanamycin and resistant to sucrose were used as a template for colony PCR using primers flanking the gene to be deleted. The primers used for each gene were the external primers used previously to amplify the flanking regions of the deletion construct. The PCR products were analyzed using gel electrophoresis on a 1% agarose gel, and colonies that yielded a PCR product with the correctly sized deleted allele size were restreaked on an ALP plate. Single colonies from the restreaked plate

were checked once again with colony PCR for the deleted allele, and then picked into an overnight ALP media culture to make into freezer stocks.

References

Achenbach, L., Michaelidou, U., Bruce, R., Fryman, J., and Coates, J. (2001) Dechloromonas agitata gen. nov., sp. nov. and Dechlorosoma suillum gen. nov., sp. nov., two novel environmentally dominant (per)chlorate-reducing bacteria and their phylogenetic position. *Int J Syst Evol Microbiol* **51**: 527–533.

Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990) Basic local alignment search tool. *Journal of molecular biology* **215**: 403–410.

Alvarez-Martinez, C.E., Baldini, R.L., and Gomes, S.L. (2006) A caulobacter crescentus extracytoplasmic function sigma factor mediating the response to oxidative stress in stationary phase. *J Bacteriol* **188**: 1835–1846.

Bender, K., Shang, C., Chakraborty, R., Belchik, S., Coates, J., and Achenbach, L. (2005) Identification, characterization, and classification of genes encoding perchlorate reductase. *J Bacteriol* **187**: 5090–5096.

Bender, K.S., O'Connor, S.M., Chakraborty, R., Coates, J.D., and Achenbach, L.A. (2002) Sequencing and transcriptional analysis of the chlorite dismutase gene of Dechloromonas agitata and its use as a metabolic probe. *Appl Environ Microbiol* **68**: 4820–4826.

Biegel, E., Schmidt, S., González, J.M., and Müller, V. (2011) Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. *Cellular and molecular life sciences : CMLS* **68**: 613–634.

Blasco, F., Santos, Dos, J.P., Magalon, A., Frixon, C., Guigliarelli, B., Santini, C.L., and Giordano, G. (1998) NarJ is a specific chaperone required for molybdenum cofactor assembly in nitrate reductase A of Escherichia coli. *Mol Microbiol* **28**: 435–447.

Bouhenni, R., Gehrke, A., and Saffarini, D. (2005) Identification of genes involved in cytochrome c biogenesis in Shewanella oneidensis, using a modified mariner transposon. *Appl Environ Microbiol* **71**: 4935–4937.

Bruce, R.A., Achenbach, L.A., and Coates, J.D. (1999) Reduction of (per)chlorate by a novel organism isolated from paper mill waste. *Environ Microbiol* **1**: 319–329.

Buelow, D.R. and Raivio, T.L. (2010) Three (and more) component regulatory systems - auxiliary regulators of bacterial histidine kinases. *Mol Microbiol* **75**: 547–566.

Bush, M. and Dixon, R. (2012) The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription. *Microbiol Mol Biol Rev* **76**: 497–529.

Byrne-Bailey, K.G. and Coates, J.D. (2012) Complete genome sequence of the anaerobic perchlorate-reducing bacterium Azospira suillum strain PS. *J Bacteriol* **194**: 2767–

2768.
Carlström, C.I., Wang, O., Melnyk, R.A., Bauer, S., Lee, J., Engelbrektson, A., and Coates, J.D. (2013) Physiological and Genetic Description of Dissimilatory Perchlorate Reduction by the Novel Marine Bacterium Arcobacter sp. Strain CAB. *mBio* **4**:.
Chaudhuri, S.K., O'Connor, S.M., Gustavson, R.L., Achenbach, L.A., and Coates, J.D. (2002) Environmental factors that control microbial perchlorate reduction. *Appl Environ Microbiol* **68**: 4425–4430.
Clark, I.C., Carlson, H.K., Iavarone, A.T., and Coates, J.D. (2012) Bioelectrical redox cycling of anthraquinone-2,6-disulfonate coupled to perchlorate reduction. *Energy Environ Sci* **5**: 7970–7978.
Clark, I.C., Melnyk, R.A., Engelbrektson, A., and Coates, J.D. (2013) Structure and evolution of chlorate reduction composite transposons. *mBio* **4**:.
Coates, J. and Achenbach, L. (2004) Microbial perchlorate reduction: rocket-fueled metabolism. *Nature Reviews Microbiology* **2**: 569–580.
Coates, J.D., Michaelidou, U., Bruce, R.A., O'Connor, S.M., Crespi, J.N., and Achenbach, L.A. (1999) Ubiquity and diversity of dissimilatory (per)chlorate-reducing bacteria. *Appl Environ Microbiol* **65**: 5234–5241.
Das, S., Noe, J.C., Paik, S., and Kitten, T. (2005) An improved arbitrary primed PCR method for rapid characterization of transposon insertion sites. *J. Microbiol. Methods* **63**: 89–94.
Galperin, M.Y. (2006) Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J Bacteriol* **188**: 4169–4182.
Galperin, M.Y., Nikolskaya, A.N., and Koonin, E.V. (2001) Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol Lett* **203**: 11–21.
Gulmezian, M., Hyman, K.R., Marbois, B.N., Clarke, C.F., and Javor, G.T. (2007) The role of UbiX in Escherichia coli coenzyme Q biosynthesis. *Arch. Biochem. Biophys.* **467**: 144–153.
Jack, R.L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T., and Sargent, F. (2004) Coordinating assembly and export of complex bacterial proteins. *EMBO J.* **23**: 3962–3972.
Jacobs, M.A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., et al. (2003) Comprehensive transposon mutant library of Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences* **100**: 14339–14344.
Jüngst, A., Wakabayashi, S., Matsubara, H., and Zumft, W. (1991) The nirSTBM region coding for cytochrome cd1-dependent nitrite respiration of Pseudomonas stutzeri consists of a cluster of mono-, di-, and tetraheme proteins. *FEBS letters* **279**: 205–209.
Kengen, S.W.S., Rikken, G.B.G., Hagen, W.R.W., van Ginkel, C.G.C., and Stams, A.J.A.

(1999) Purification and characterization of (per)chlorate reductase from the chlorate-respiring strain GR-1. *J Bacteriol* **181**: 6706–6711.
Kloer, D.P., Hagel, C., Heider, J., and Schulz, G.E. (2006) Crystal structure of ethylbenzene dehydrogenase from Aromatoleum aromaticum. *Structure* **14**: 1377–1388.
Kohler, C., Lourenço, R.F., Avelar, G.M., and Gomes, S.L. (2012) Extracytoplasmic function (ECF) sigma factor sigmaF is involved in Caulobacter crescentus response to heavy metal stress. *BMC Microbiol* **12**: 210.
Kounaves, S.P., Stroble, S.T., Anderson, R.M., Moore, Q., Catling, D.C., Douglas, S., et al. (2010) Discovery of natural perchlorate in the Antarctic Dry Valleys and its global implications. *Environ Sci Technol* **44**: 2360–2364.
Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175–176.
Kwolek-Mirek, M., Bartosz, G., and Spickett, C.M. (2011) Sensitivity of antioxidant-deficient yeast to hypochlorite and chlorite. *Yeast* **28**: 595–609.
Lam, C.K., O'Mullan, P., and Eveleigh, D.E. (1993) Transformation of Zymononas mobilis by electroporation. *Applied microbiology and Biotechnology* **39**: 305–308.
Larsen, R.A., Wilson, M.M., Guss, A.M., and Metcalf, W.W. (2002) Genetic analysis of pigment biosynthesis in Xanthobacter autotrophicus Py2 using a new, highly efficient transposon mutagenesis system that is functional in a wide variety of bacteria. *Arch Microbiol* **178**: 193–201.
Lee, A.I., Delgado, A., and Gunsalus, R.P. (1999) Signal-dependent phosphorylation of the membrane-bound NarX two-component sensor-transmitter protein of Escherichia coli: nitrate elicits a superior anion ligand response compared to nitrite. *J Bacteriol* **181**: 5309–5316.
Loschi, L., Brokx, S.J., Hills, T.L., Zhang, G., Bertero, M.G., Lovering, A.L., et al. (2004) Structural and biochemical identification of a novel bacterial oxidoreductase. *The Journal of biological chemistry* **279**: 50391–50400.
Maillard, J., Spronk, C.A.E.M., Buchanan, G., Lyall, V., Richardson, D.J., Palmer, T., et al. (2007) Structural diversity in twin-arginine signal peptide-binding proteins. *Proceedings of the National Academy of Sciences* **104**: 15641–15646.
Melnyk, R.A., Engelbrektson, A., Clark, I.C., Carlson, H.K., Byrne-Bailey, K., and Coates, J.D. (2011) Identification of a perchlorate reduction genomic island with novel regulatory and metabolic genes. *Appl Environ Microbiol* **77**: 7401–7404.
Mlynek, G., Sjöblom, B., Kostan, J., Füreder, S., Maixner, F., Gysel, K., et al. (2011) Unexpected Diversity of Chlorite Dismutases: a Catalytically Efficient Dimeric Enzyme from Nitrobacter winogradskyi. *J Bacteriol* **193**: 2408–2417.
Möglich, A., Ayers, R.A., and Moffat, K. (2009) Structure and signaling mechanism of

Per-ARNT-Sim domains. *Structure* **17**: 1282–1294.
Murat, D., Quinlan, A., Vali, H., and Komeili, A. (2010) Comprehensive genetic dissection of the magnetosome gene island reveals the step-wise assembly of a prokaryotic organelle. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 5593–5598.
Nadaraja, A.V., Gangadharan Puthiya Veetil, P., and Bhaskaran, K. (2012) Perchlorate reduction by an isolated *Serratia marcescens* strain under high salt and extreme pH. *FEMS Microbiol Lett* **339**: 117–121.
Parker, D.R. (2009) Perchlorate in the environment: the emerging emphasis on natural occurrence. *Environmental Chemistry* **6**: 10–27.
Rajagopalan, S., Anderson, T., Cox, S., Harvey, G., Cheng, Q., and Jackson, W.A. (2009) Perchlorate in wet deposition across North America. *Environ Sci Technol* **43**: 616–622.
Reinhold-Hurek, B. and Hurek, T. (2006) The Genera Azoarcus, Azovibrio, Azospira and Azonexus. Springer New York, New York, NY, pp. 873–891.
Schwarz, G., Mendel, R.R., and Ribbe, M.W. (2009) Molybdenum cofactors, enzymes and pathways. *Nature* **460**: 839–847.
Shi, L., Squier, T., Zachara, J., and Fredrickson, J. (2007) Respiration of metal (hydr)oxides by Shewanella and Geobacter: a key role for multihaem c-type cytochromes. *Mol Microbiol* **65**: 12–20.
Simon, J. and Kern, M. (2008) Quinone-reactive proteins devoid of haem b form widespread membrane-bound electron transport modules in bacterial respiration. *Biochemical Society transactions* **36**: 1011–1016.
Stenklo, K., Thorell, H.D., Bergius, H., Aasa, R., and Nilsson, T. (2001) Chlorite dismutase from Ideonella dechloratans. *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry* **6**: 601–607.
Sun, Y., Gustavson, R., Ali, N., Weber, K., Westphal, L., and Coates, J. (2009) Behavioral response of dissimilatory perchlorate-reducing bacteria to different electron acceptors. *Appl Microbiol Biotechnol* **84**: 955–963.
Thorell, H.D., Karlsson, J., Portelius, E., and Nilsson, T. (2002) Cloning, characterisation, and expression of a novel gene encoding chlorite dismutase from Ideonella dechloratans. *Biochim Biophys Acta* **1577**: 445–451.
Thrash, J.C., Ahmadi, S., Torok, T., and Coates, J.D. (2010) Magnetospirillum bellicus sp. nov., a novel dissimilatory perchlorate-reducing alphaproteobacterium isolated from a bioelectrical reactor. *Appl Environ Microbiol* **76**: 4730–4737.
Thrash, J.C., Pollock, J., Torok, T., and Coates, J.D. (2010) Description of the novel perchlorate-reducing bacteria Dechlorobacter hydrogenophilus gen. nov., sp. nov.and Propionivibrio militaris, sp. nov. *Appl Microbiol Biotechnol* **86**: 335–343.
Troup, B., Hungerer, C., and Jahn, D. (1995) Cloning and characterization of the Escherichia coli hemN gene encoding the oxygen-independent coproporphyrinogen

III oxidase. *J Bacteriol* **177**: 3326–3331.
van Ginkel, C.G., Rikken, G.B., Kroon, A.G.M., and Kengen, S.W.M. (1996) Purification and characterization of chlorite dismutase: a novel oxygen-generating enzyme. *Arch Microbiol* **166**: 321–326.
Van Trump, J.I. and Coates, J.D. (2009) Thermodynamic targeting of microbial perchlorate reduction by selective electron donors. *Isme J* **3**: 466–476.
Xu, J., Trimble, J., Steinberg, L., and Logan, B. (2004) Chlorate and nitrate reduction pathways are separately induced in the perchlorate-respiring bacterium Dechlorosoma sp. KJ and the chlorate-respiring bacterium Pseudomonas sp. PDA. *Water research* **38**: 673–680.
Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T.L. (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics* **13**: 134.

Tables and Figures

Figure 1. *PS growth on nitrate and perchlorate.* Wild-type PS was grown on media containing 5 mM nitrate, 2.5 mM perchlorate, and a mixture of nitrate and perchlorate (5 mM each). This experiment was carried out on rich ALP media.

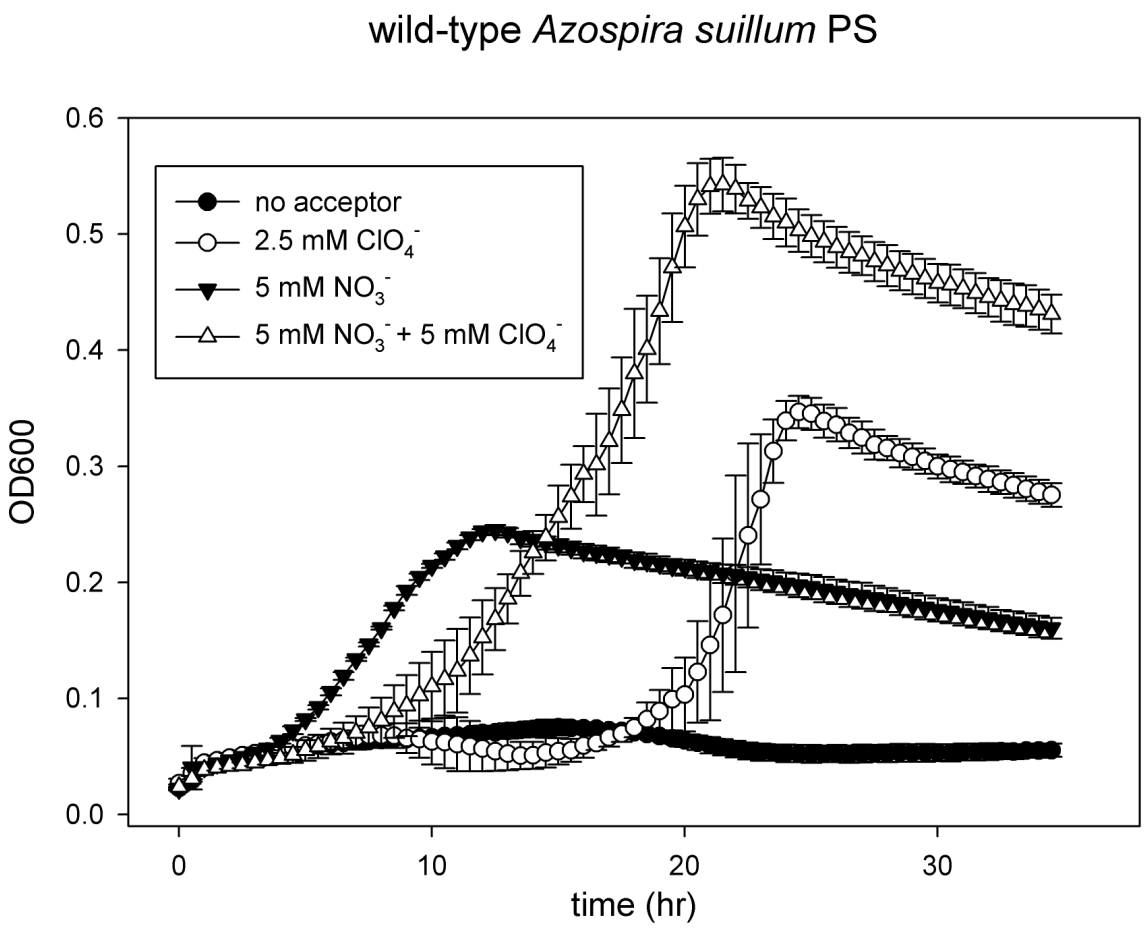


Figure 2. *Location of transposon insertions in the PRI.* The 17 genes in the PRI are displayed with the locations of transposon insertions that were identified in the screen. The green pins indicate mutants that did not grow on perchlorate, while the blue pins indicates mutants that did not grow on perchlorate or on the diauxic media. Genes are colored by functional group: green genes encode chlorite dismutase and perchlorate reductase components, royal blue genes are the PRI HKS, red genes are parts of the putative electron transport chain, orange genes are components of a putative oxidoreductase system, and light blue genes are part of a putative sigma factor/anti-sigma system.

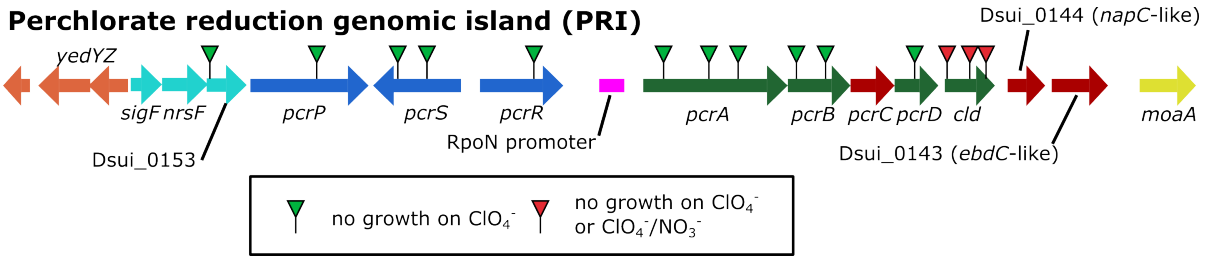


Figure 3. *Growth of pcrA::Himar and cld::Himar.* The growth of strains with transposons on A) nitrate, B) perchlorate, and C) under diauxic conditions.

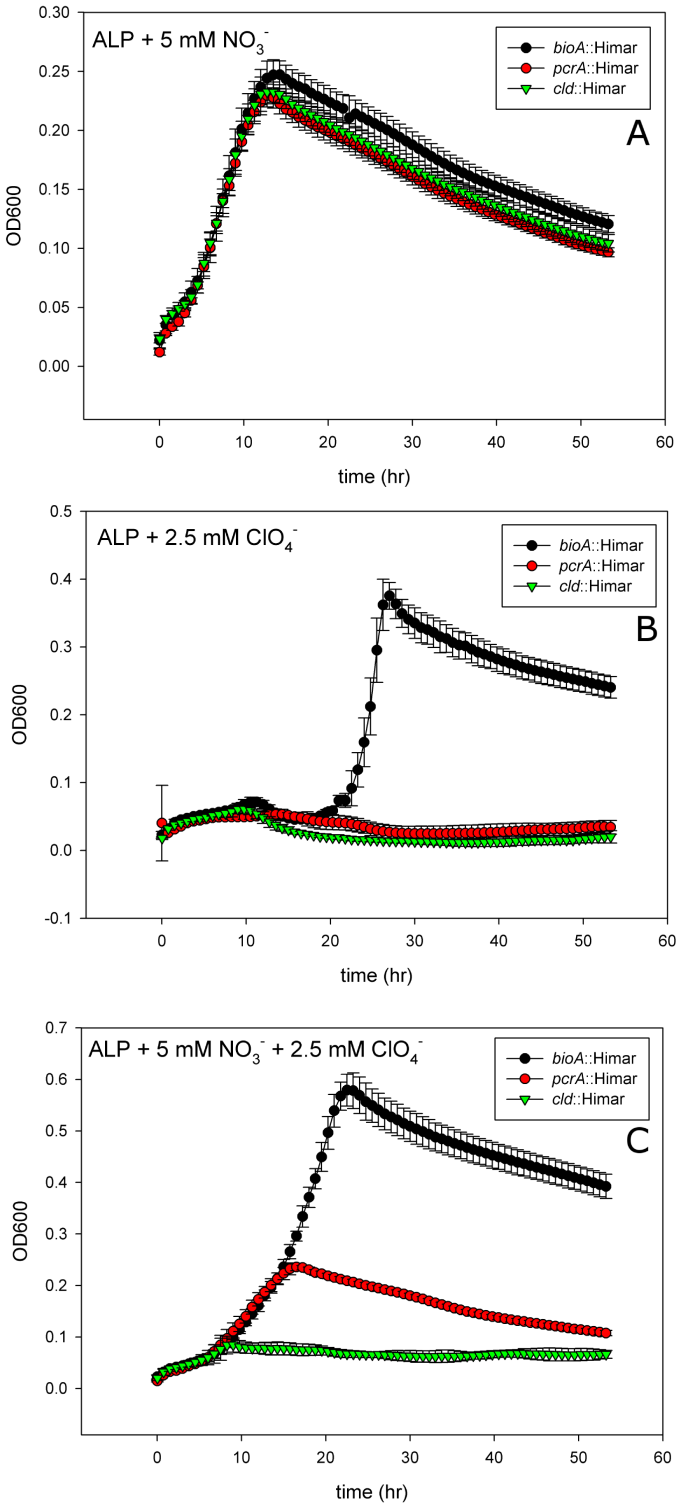


Figure 4. *Complementation of ΔpcrP, ΔpcrS, and ΔpcrR.* The three knockout strains of the elements of the histidine kinase system, as well as wild-type PS were transformed either with a complementation vector (pRAMXX) or an empty vector (pBBR1MCS2) prior to growth on diauxic ALP media.

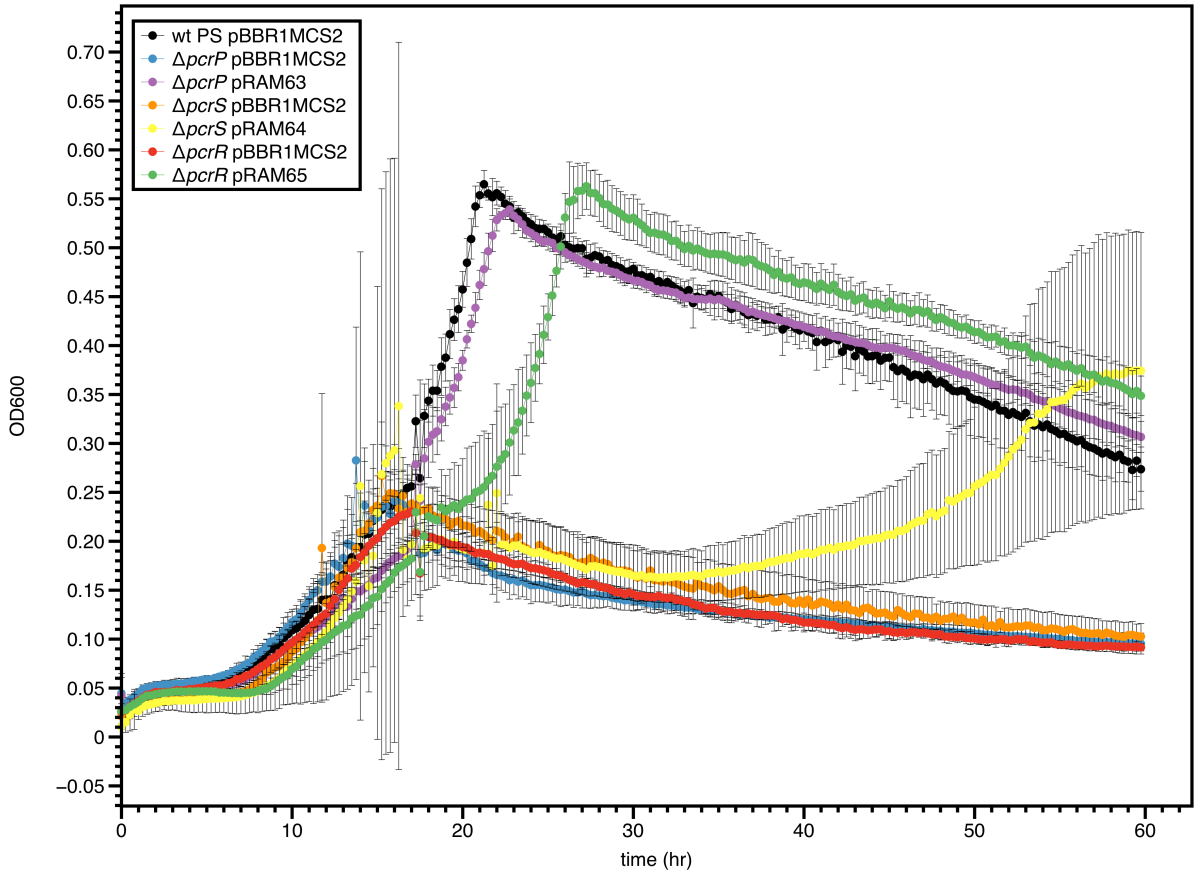


Figure 5. *Complementation of ΔpcrA and ΔpcrC.* The knockout strains of Δ*pcrA* and Δ*pcrC*, as well as wild-type PS, were transformed either with a complementation vector (pRAMXX) or an empty vector (pBBR1MCS2) prior to growth on diauxic ALP media.

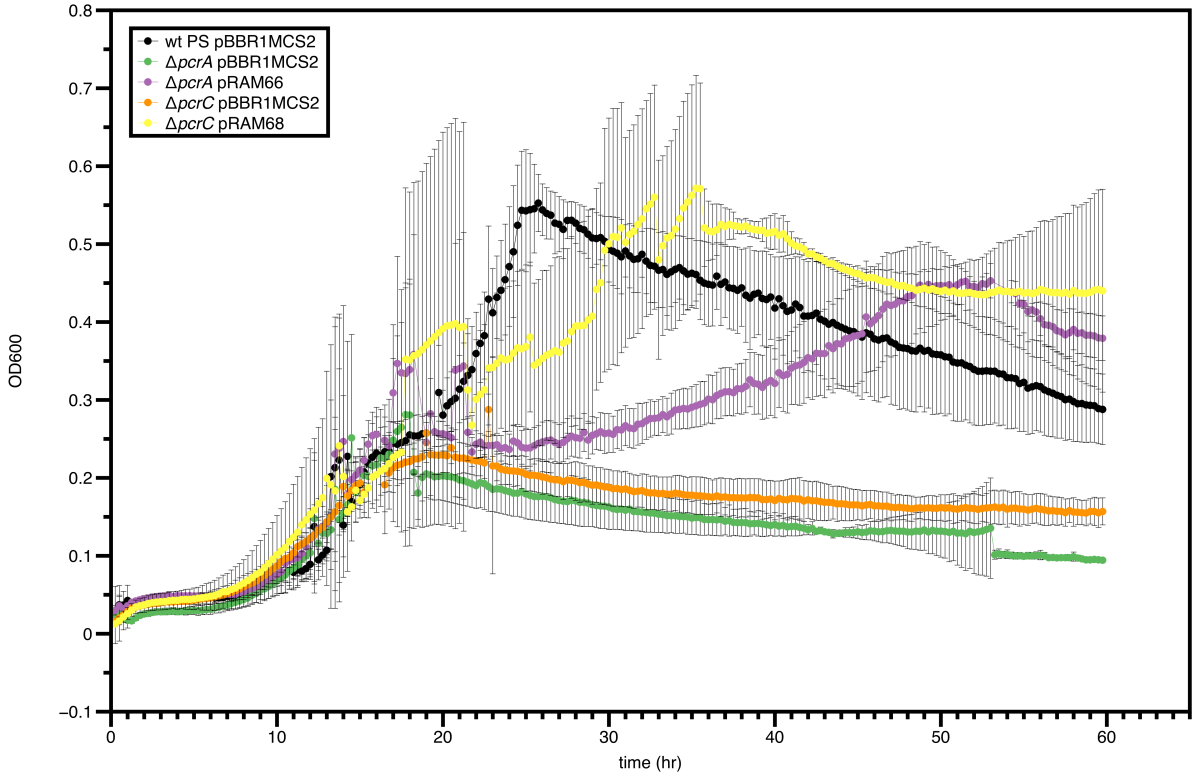


Figure 6. *Complementation of ΔpcrB, ΔpcrD, and Δcld.* The knockout strains of Δ*pcrB*, Δ*pcrD* and Δ*cld*, as well as wild-type PS, were transformed either with a complementation vector (pRAMXX) or an empty vector (pBBR1MCS2) prior to growth on diauxic ALP media.

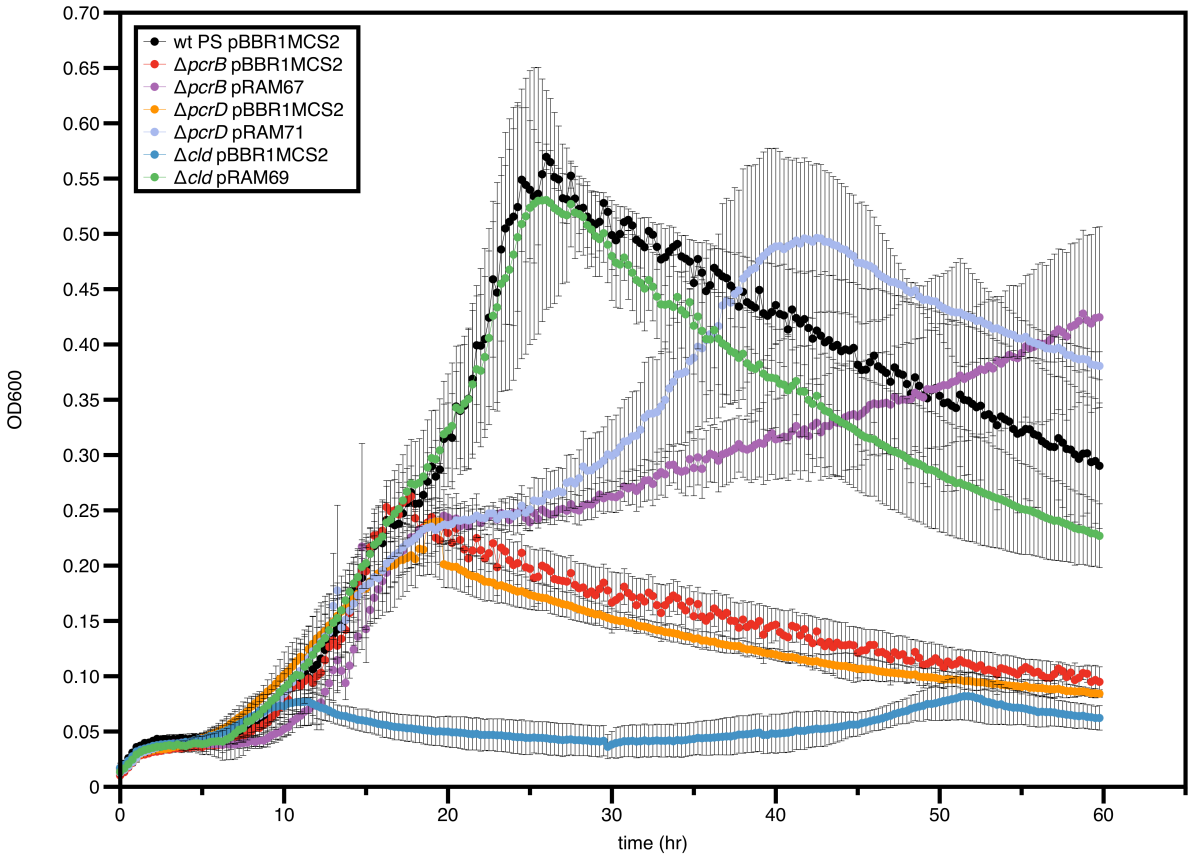


Figure 7. *Growth characteristics of wt PS, Δcld, ΔpcrA and ΔcldΔpcrA strains.* Growth on A) nitrate, B) diauxic conditions, and C) perchlorate.

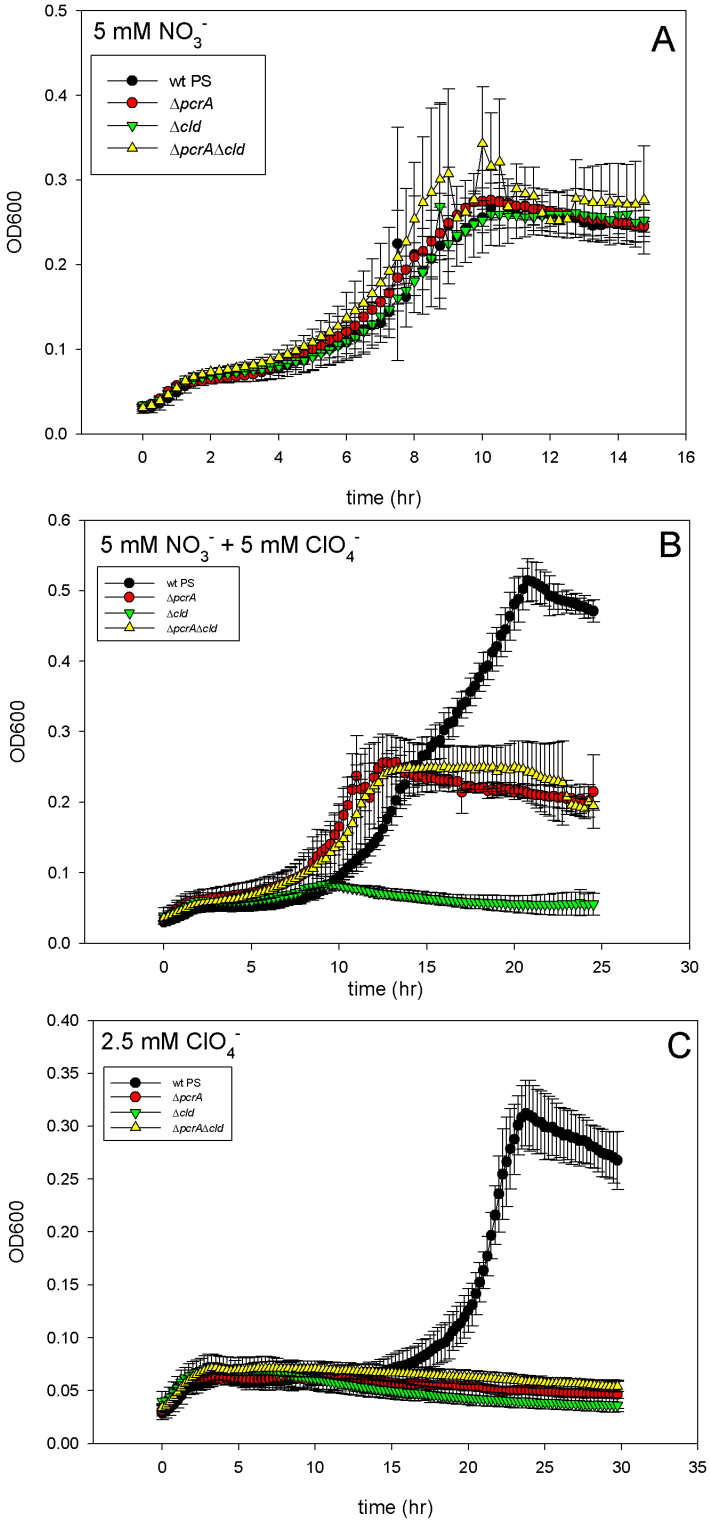


Figure 8. *A model of perchlorate reduction in Azospira suillum PS.* Gene products from the PRI are depicted in a model that reflects our current understanding of perchlorate reduction and its regulation in *Azospira suillum* PS. Several genes identified in the transposon screen as important for perchlorate reduction are also shown (molybdopterin cofactor biosynthesis, denitrification, and the Rnf complex). Gene products from the PRI are colored to be consistent with the functional group coloring in Figure 2.

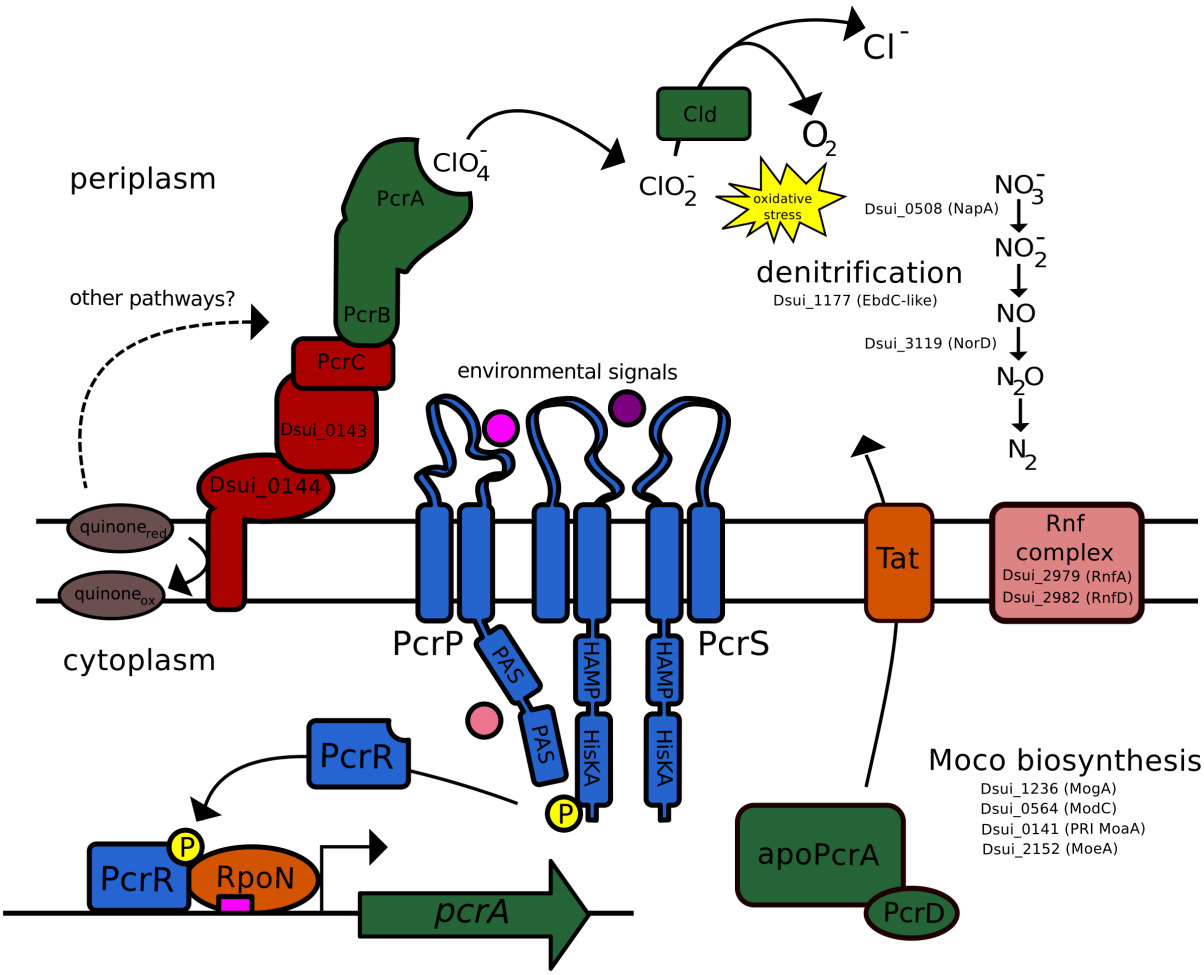


Table 1. *Transposon mutants mapped and characterized in this chapter.* Transposon mutants were scored "++" if growth was identical to wild-type, "+" if growth rate or final optical density were less than wild-type, or "0" if no growth was observed. "N" indicates growth on 5 mM nitrate, "P" indicates growth on 2.5 mM perchlorate, and "N/P" indicates diauxic growth on nitrate and perchlorate (5 mM each).

Mutant ID	Locus Tag	Gene Name and Gene Product	Phenotype	N	P	N/P	Aerobic
A3C3	Dsui_1578	*aceE* - pyruvate dehydrogenase	pleiotropic	+	+	+	+
A3G1	Dsui_0149	*pcrA* - perchlorate reductase alpha subunit	perchlorate null	++	0	+	++
A4A8	Dsui_0150	*pcrR* - PRI response regulator	perchlorate null	++	0	+	++
A4A9	Dsui_0151	*pcrS* - PRI histidine kinase	perchlorate null	++	0	+	++
A5B2	Dsui_2880	*bioA* - adenosylmethionine--8-amino-7-oxononanoate transaminase	no phenotype	++	++	++	++
A2D6	Dsui_0261	*cysP* - periplasmic binding protein of sulfate/thiosulfate ABC transporter	pleiotropic	+	+	+	++
A4F1	Dsui_0145	*cld* - chlorite dismutase	perchlorate null	++	0	0	++
A5A6	Dsui_0704	*rpoN* - sigma-54 type sigma factor RpoN	perchlorate null	++	0	+	++
A5D9	Dsui_0149	*pcrA* - perchlorate reductase alpha subunit	perchlorate null	++	0	+	++
A7B1	Dsui_1177	heme *b*/*c*-containing EbdC-like electron transport protein	pleiotropic	0	+	+	++
A7B10	Dsui_0149	*pcrA* - perchlorate reductase alpha subunit	perchlorate null	++	0	+	++
A7E8	Dsui_1441	phenylacetate-CoA ligase	perchlorate null	++	0	+	++
A8E1	Dsui_0508	*napA* - nitrate reductase alpha subunit	pleiotropic	+	0	0	++
A8H2	Dsui_0675	*dsbA* - thiol:disulfide interchange protein	pleiotropic	+	+	+	++
A9A5	Dsui_0153	PRI cupin domain protein	perchlorate null	++	0	+	++
A9G11	Dsui_0993	*ubiX* - 3-octaprenyl-4-hydroxybenzoate carboxylyase	pleiotropic	+	+	+	++
B1H6	Dsui_0145	*cld* - chlorite dismutase	perchlorate null	++	0	0	++
B1F10	Dsui_2982	*rnfD* - Rnf electron transport complex delta subunit	pleiotropic	++	+	+	++
B2D12	Dsui_3119	*norD* - nitric oxide reductase activation protein	pleiotropic	+	++	+	++
B2H11	Dsui_1236	*mogA* – pterin adenyltransferase	pleiotropic	+	0	+	++

B3E6	Dsui_1533	*dsbB* - disulfide bond formation protein	pleiotropic	+	++	+	++
B4D1	Dsui_0704	*rpoN* - sigma-54 type sigma factor RpoN	perchlorate null	++	0	+	++
B4E5	downstream of Dsui_0508	*napA* - nitrate reductase alpha subunit	pleiotropic	+	+	+	++
B6F5	Dsui_0564	*modC* - ATPase subunit of molybdate ABC transporter	pleiotropic	++	+	+	++
B7F10	Dsui_1054	*trpB* - tryptophan synthase	pleiotropic	++	+	+	++
B7G4	Dsui_0151	*pcrS* - PRI histidine kinase	perchlorate null	++	0	+	++
B9B5	Dsui_2152	*moeA* - molybdopterin molybdotransferase	pleiotropic	+	+	+	++
B9B7	Dsui_0146	*pcrD* - perchlorate reductase cytoplasmic chaperone	perchlorate null	++	0	+	++
B9G5	Dsui_0148	*pcrB* - perchlorate reductase beta subunit	perchlorate null	++	0	+	++
B9H11	Dsui_2979	*rnfA* - Rnf electron transport complex alpha subunit	pleiotropic	+	+	+	++
C1A4	Dsui_1647	*hemN* - oxygen-independent coproporphyrinogen III oxidase	pleiotropic	+	+	+	++
C1D2	Dsui_0152	*pcrP* - PRI PAS domain protein	perchlorate null	++	0	+	++
C4E9	Dsui_2816	(p)ppGpp synthetase, RelA/SpoT family	pleiotropic	++	+	+	++
C4H3	Dsui_0128	D-alanyl-D-alanine carboxypeptidase	perchlorate null	++	0	+	++
C5A2	Dsui_0148	*pcrB* - perchlorate reductase beta subunit	perchlorate null	++	0	+	++
C5B11	Dsui_0145	*cld* - chlorite dismutase	perchlorate null	++	0	0	++
C5B9	upstream of Dsui_2979	*rnfA* - Rnf electron transport complex alpha subunit	pleiotropic	++	+	+	++

Table 2. *Annotations of the 17 core PRI genes and phenotypic information based on deletion mutants.* For growth on perchlorate, "++" denotes wild-type growth on perchlorate, "+" denotes growth at a slower rate than wild-type, and "0" denotes no growth. For growth on nitrate and perchlorate, "++" denotes wild-type growth, "+" denotes growth only during the initial denitrification phase, and "0" denotes very little growth. All deletions grew identically to wild-type on nitrate alone.

Locus Tag	Gene Name	Gene Product	Pfam Domains	P	N/P
Dsui_0141	*moaA*	molybdenum cofactor biosynthesis protein	PF04055, PF06463, PF13353	+ +	++
Dsui_0143	-	heme *b*/*c*-containing EbdC-like electron transport protein	PF09459	+	++
Dsui_0144	-	quinol dehydrogenase tetraheme cytochome *c* (NapC-like)	PF03264	+	++
Dsui_0145	*cld*	chlorite dismutase	PF06778	0	0
Dsui_0146	*pcrD*	perchlorate reductase cytoplasmic chaperone	PF02613	0	+
Dsui_0147	*pcrC*	tetraheme cytochome *c*	PF13435	0	+
Dsui_0148	*pcrB*	perchlorate reductase beta subunit	PF13247	0	+
Dsui_0149	*pcrA*	perchlorate reductase alpha subunit	PF00384, PF01568	0	+
Dsui_0150	*pcrR*	PRI response regulator	PF00072, PF00158, PF02954	0	+
Dsui_0151	*pcrS*	PRI sensor histidine kinase	PF00512, PF00672, PF02518	0	+
Dsui_0152	*pcrP*	PRI PAS domain protein	PF08448, PF13188	0	+
Dsui_0153	-	cupin domain protein	PF07883	+ +	++
Dsui_0154	*nrsF*	anti-sigma factor	PF06532	+ +	++
Dsui_0155	*sigF*	ECF sigma factor	PF04542, PF08281	+ +	++
Dsui_0156	*yedZ*	cytochrome *b* membrane electron transport protein	PF00033	+ +	++
Dsui_0157	*yedY*	unknown molybdopterin oxidoreductase	PF00174	+ +	++
Dsui_0158	-	hypothetical protein	-	+ +	++

Table 3. *Strains used in this chapter.*

Strain Name	Species	Plasmid	Genotype (description)	Source or Reference
PS	*Azospira suillum* PS	none	wild-type	Coates Lab, University of California, Berkeley (Achenbach *et al.*, 2001)
WM3064	*Escherichia coli*	none	*thrB*1004 *pro thi rpsL hsdS lac*ZDM15 RP4-1360 D(*araBAD)567* D*dapA*1341::[*erm pir*(wt)]	W. Metcalf, University of Illinois via Arash Komeili (Larsen *et al.*, 2002)
APA30	*Escherichia coli*	pMiniHimarRB1	*thrB*1004 *pro thi rpsL hsdS lac*ZDM15 RP4-1360 D(*araBAD)567* D*dapA*1341::[*erm pir*(wt)]	Arkin Lab, University of California, Berkeley
RAM22	*Azospira suillum* PS	none	Δ*pcrR*	This work
RAM24	*Azospira suillum* PS	none	Δ*pcrP*	This work
RAM25	*Azospira suillum* PS	none	Δ*pcrS*	This work
RAM26	*Azospira suillum* PS	none	Δ*nrsF*	This work
RAM27	*Azospira suillum* PS	none	Δ*yedY*	This work
RAM28	*Azospira suillum* PS	none	Δ*sigF*	This work
RAM29	*Azospira suillum* PS	none	Δ*yedZ*	This work
RAM30	*Azospira suillum* PS	none	ΔDsui_0158	This work
RAM31	*Azospira suillum* PS	pBBR1MCS2	wild-type	This work
RAM33	*Azospira suillum* PS	pBBR1MCS2	Δ*pcrR*	This work
RAM35	*Azospira suillum* PS	none	Δ*pcrA*	This work
RAM36	*Azospira suillum* PS	none	Δ*pcrB*	This work
RAM38	*Azospira suillum* PS	none	ΔDsui_0143	This work
RAM39	*Azospira suillum* PS	none	Δ*pcrD*	This work
RAM40	*Azospira suillum* PS	none	ΔDsui_0153	This work
RAM42	*Azospira suillum* PS	none	Δ*cld*	This work
RAM43	*Azospira suillum* PS	none	Δ*cld*Δ*pcrA*	This work
RAM44	*Azospira suillum* PS	none	ΔDsui_0141	This work
RAM45	*Azospira suillum* PS	none	Δ*pcrC*	This work
RAM51	*Azospira suillum* PS	none	ΔDsui_0144	This work
RAM58	*Azospira suillum* PS	pBBR1MCS2	Δ*pcrP*	This work
RAM59	*Azospira suillum* PS	pBBR1MCS2	Δ*pcrS*	This work
RAM60	*Azospira suillum* PS	pRAM64	Δ*pcrS*	This work
RAM61	*Azospira suillum* PS	pRAM65	Δ*pcrR*	This work
RAM62	*Azospira suillum* PS	pRAM63	Δ*pcrP*	This work
RAM63	*Azospira suillum* PS	pBBR1MCS2	Δ*pcrA*	This work
RAM64	*Azospira suillum* PS	pRAM66	Δ*pcrA*	This work
RAM67	*Azospira suillum* PS	pBBR1MCS2	Δ*pcrC*	This work

RAM68	*Azospira suillum* PS	pRAM68	Δ*pcrC*	This work
RAM69	*Azospira suillum* PS	pBBR1MCS2	Δ*pcrB*	This work
RAM70	*Azospira suillum* PS	pBBR1MCS2	Δ*cld*	This work
RAM71	*Azospira suillum* PS	pRAM69	Δ*cld*	This work
RAM72	*Azospira suillum* PS	pRAM67	Δ*pcrB*	This work
RAM73	*Azospira suillum* PS	pBBR1MCS2	Δ*pcrD*	This work
RAM74	*Azospira suillum* PS	pRAM71	Δ*pcrD*	This work

Table 4. *Plasmids used in this chapter.*

Plasmid	Markers	Description	Source or Reference
pNPTS138	Kan^R, Suc^S	mobilizable suicide vector	Dickon Alley (via Kathleen Ryan)
pBBR1MCS2	Kan^R	5.0-kb broad-host-range vector for cloning and complementation	Kovach *et al.*, 1995
pRAM21	Kan^R, Suc^S	pNPTS138 with Dsui_0150 deletion insert	This work
pRAM25	Kan^R, Suc^S	pNPTS138 with Dsui_0152 deletion insert	This work
pRAM26	Kan^R, Suc^S	pNPTS138 with Dsui_0151 deletion insert	This work
pRAM33	Kan^R, Suc^S	pNPTS138 with Dsui_0156 deletion insert	This work
pRAM37	Kan^R, Suc^S	pNPTS138 with Dsui_0157 deletion insert	This work
pRAM38	Kan^R, Suc^S	pNPTS138 with Dsui_0154 deletion insert	This work
pRAM39	Kan^R, Suc^S	pNPTS138 with Dsui_0155 deletion insert	This work
pRAM40	Kan^R, Suc^S	pNPTS138 with Dsui_0158 deletion insert	This work
pRAM44	Kan^R, Suc^S	pNPTS138 with Dsui_0149 deletion insert	This work
pRAM45	Kan^R, Suc^S	pNPTS138 with Dsui_0148 deletion insert	This work
pRAM48	Kan^R, Suc^S	pNPTS138 with Dsui_0145 deletion insert	This work
pRAM49	Kan^R, Suc^S	pNPTS138 with Dsui_0146 deletion insert	This work
pRAM50	Kan^R, Suc^S	pNPTS138 with Dsui_0141 deletion insert	This work
pRAM51	Kan^R, Suc^S	pNPTS138 with Dsui_0143 deletion insert	This work
pRAM52	Kan^R, Suc^S	pNPTS138 with Dsui_0153 deletion insert	This work
pRAM54	Kan^R, Suc^S	pNPTS138 with Dsui_0147 deletion insert	This work
pRAM55	Kan^R, Suc^S	pNPTS138 with Dsui_0144 deletion insert	This work
pRAM62	Kan^R	pBBR1MCS2 with *pcrA* promoter	This work
pRAM63	Kan^R	pBBR1MCS2 with *pcrP* promoter/gene	This work
pRAM64	Kan^R	pBBR1MCS2 with *pcrS* promoter/gene	This work
pRAM65	Kan^R	pBBR1MCS2 with *pcrR* promoter/gene	This work
pRAM66	Kan^R	pRAM62 with *pcrA* gene	This work
pRAM67	Kan^R	pRAM62 with *pcrB* gene	This work
pRAM68	Kan^R	pRAM62 with *pcrC* gene	This work
pRAM69	Kan^R	pRAM62 with *cld* gene	This work
pRAM71	Kan^R	pRAM62 with *pcrD* gene	This work
pMiniHimar RB1	Kan^R	suicide vector for transposon delivery	Bouhenni *et al.*, 2005

Table 5. *Primers used in this chapter.*

Primer ID	Primer Sequence	Brief Description of Primer
HIMAR_EXT	CCGTGATATTGCTGAAGAGC	Arbitrary PCR Tn mapping Round 1
HIMAR_INT	TGACGAGTTCTTCTGAGCGG	Arbitrary PCR Tn mapping Round 2 & Sanger sequencing
PS_ARB2	GGCCACGCGTCGACTAGTAC	Arbitrary PCR Tn mapping Round 2
PS_ARB4	GGCCACGCGTCGACTAGTACNNNNNNNNNNGCCAG	Arbitrary PCR Tn mapping Round 1
PS_ARB5	GGCCACGCGTCGACTAGTACNNNNNNNNNNGGCGG	Arbitrary PCR Tn mapping Round 1
PS_ARB6	GGCCACGCGTCGACTAGTACNNNNNNNNNNCGGCG	Arbitrary PCR Tn mapping Round 1
PS0141a	GNNGAATTCGACAAAGTTTCGGCTGCCTC	deletion of Dsui_0141
PS0141b	GNNACTAGTTATCCGACGACCACAACCATC	deletion of Dsui_0141
PS0141c	GNNACTAGTGGATAATTAGCGCTCTGGCG	deletion of Dsui_0141
PS0141d	GNNGGATCCGGCTTGATCCCCAAAAGGGA	deletion of Dsui_0141
PS0143a	GNNGGATCCCGGAAGATGTGATCAAGGCGCTAGCCGATTA	deletion of Dsui_0143
PS0143b	GNNACTAGTACCCAAAATAGCTACTGATAAAGCA	deletion of Dsui_0143
PS0143c	GNNACTAGTGGCGCTGATGGCGATATTAA	deletion of Dsui_0143
PS0143d	GNNGAATTCAATCCACTCTCTCGTGCGAC	deletion of Dsui_0143
PS0144a	GNNGGATCCTTACGCCCGGATTGGAACAA	deletion of Dsui_0144
PS0144b	GNNACTAGTGCGCAGCGATTTGAGAATGT	deletion of Dsui_0144
PS0144c	GNNACTAGTGACGAGTGATCGTCCTCGCAT	deletion of Dsui_0144
PS0144d	GNNGAATTCCATCAGCGCCAAAACCAAGT	deletion of Dsui_0144
PS0145a	GNNGAATTCCGCAACGCTTCAGGAAGATG	deletion of Dsui_0145
PS0145b	GNNGGATCCGCGAATTGATATGTTTGTCATGG	deletion of Dsui_0145
PS0145c	GNNGGATCCATCAAGGCGCTAGCCGATTA	deletion of Dsui_0145
PS0145d	GNNACTAGTGCGATTTGAGACCCAAGGGA	deletion of Dsui_0145
PS0146a	GNNGAATTCTAATGTTGCAATGGAGCGGC	deletion of Dsui_0146
PS0146b	GNNGGATCCGTCGACAAGGCTATTCATTTG	deletion of Dsui_0146
PS0146c	GNNGGATCCATTCTGCGGGAAATTGATTCTT	deletion of Dsui_0146
PS0146d	GNNACTAGTTGTGCCAGAGAGAGCATCAG	deletion of Dsui_0146
PS0147a	GNNACTAGTTCACCGTTCTTGTAGCCACC	deletion of Dsui_0147
PS0147b	GNNGGATCCGGTGGCGAGAGCAAGAGTTT	deletion of Dsui_0147
PS0147c	GNNGGATCCCCTGAGCCCGAATAACGCTGA	deletion of Dsui_0147

PS0148a	GNNACTAGTGAGAAGTTCAAGTGGGTGCG	deletion of Dsui_0148
PS0148b	GNNGAATTCGGGCGCTTTCATCACATTAGC	deletion of Dsui_0148
PS0148c	GNNGAATTCCGCAGCACCGACATGATGATT	deletion of Dsui_0148
PS0148d	GNNGGATCCTCAGCGTTATTCGGGCTCAG	deletion of Dsui_0148
PS0149a	GNNGAATTCTAGCCGTCGTGATTGGATGG	deletion of Dsui_0149
PS0149b	GNNGGATCCGGCTAGCAGGAAGCCTCTACG	deletion of Dsui_0149
PS0149c	GNNGGATCCGCCACGCCCATCAGTTTCTA	deletion of Dsui_0149
PS0149d	GNNACTAGTCCAAGTACAGGCGGCACATA	deletion of Dsui_0149
PS0150a	GNNACTAGTCGTCTCCGGACGGGACTGGT	deletion of Dsui_0150
PS0150b	TATATAAATTAAGTCGATGGGGTGTATCAATTCACGATCTCCCGC	deletion of Dsui_0150
PS0150c	CCCATCGACTTAATTTATATAAGCCCTCCTGATCGTACAAGTGGC	deletion of Dsui_0150
PS0150d	GNNACTAGTGCCAATGGTTGGCAGCCTGA	deletion of Dsui_0150
PS0151a	GNNACTAGTCGGCGCCAGTGAAAGCACCT	deletion of Dsui_0151
PS0151b	TATATAAATTAAGTCGATGGGTGCTGCGCTCCAGGTTAGAACA	deletion of Dsui_0151
PS0151c	CCCATCGACTTAATTTATATAACGGATGACTCACCACAAAGTGGC	deletion of Dsui_0151
PS0151d	GNNACTAGTTGGAAGCGATTGCCGAAGGCT	deletion of Dsui_0151
PS0152a	GNNACTAGTGCGCTCGGCCTTCCTGTCTT	deletion of Dsui_0152
PS0152b	TATATAAATTAAGTCGATGGGTCCGGCAGTTTGCAGAGCGG	deletion of Dsui_0152
PS0152c	CCCATCGACTTAATTTATATAAAACTGCGCGATCGGGACGA	deletion of Dsui_0152
PS0152d	GNNACTAGTTCAGTCGGGCTATGACGCGGA	deletion of Dsui_0152
PS0153a	GNNGGATCCGTCAGAGAAGCGGCTGATGT	deletion of Dsui_0153
PS0153b	GNNACTAGTTCGTTTTAGTGGCACGGTAGA	deletion of Dsui_0153
PS0153c	GNNACTAGTTACTTTTTAGCCTGAGGTGCGA	deletion of Dsui_0153
PS0153d	GNNGAATTCGGGGTCGGTACGTATCACAG	deletion of Dsui_0153
PS0154a	GNNACTAGTGCCGCCGAGCGTGATTTCCT	deletion of Dsui_0154
PS0154b	TATATAAATTAAGTCGATGGGTCCCGCGGCCAACATGGTGAT	deletion of Dsui_0154
PS0154c	CCCATCGACTTAATTTATATAATGCTGCGCTGGTAGCCTTTTCG	deletion of Dsui_0154
PS0154d	GNNACTAGTACGTTCTGCAAGCCAAGCCGA	deletion of Dsui_0154
PS0155a	GNNACTAGTGTGATGGCGCACCCGGTCAG	deletion of Dsui_0155
PS0155b	TATATAAATTAAGTCGATGGGTCGTTGCATGGACGCCTTTCGT	deletion of Dsui_0155
PS0155c	CCCATCGACTTAATTTATATAGCGCTCGCGGCGAAAATACG	deletion of Dsui_0155
PS0155d	GNNACTAGTAGGTCGAATGCGTTGCCCCG	deletion of Dsui_0155
PS0156a	GNNACTAGTGCAAGAGCGCAAGCCCGAGA	deletion of Dsui_0156
PS0156b	TATATAAATTAAGTCGATGGGCTCTTTGCTCATGGCGTTCCTC	deletion of Dsui_0156
PS0156c	CCCATCGACTTAATTTATATAATCGCGATGATTCGCGGCCA	deletion of Dsui_0156
PS0156d	GNNACTAGTCCGACGGCGACCATAAGGCA	deletion of Dsui_0156
PS0157a	GNNACTAGTTTACCGGACAGGCCACGCAC	deletion of Dsui_0157
PS0157b	TATATAAATTAAGTCGATGGGGTTGGCAATCATCGGGCGGTTC	deletion of Dsui_0157
PS0157c	CCCATCGACTTAATTTATATAGGCGGCAGTTGATCGTTTCGG	deletion of Dsui_0157

PS0157d	GNNACTAGTGCCAAGCCTGCCGTTCGACT	deletion of Dsui_0157
PS0158a	GNNACTAGTTGGCCATGGTCGCGTTGGTC	deletion of Dsui_0158
PS0158b	TATATAAATTAAGTCGATGGGGGTGAGCATTACTGCGGCGAG	deletion of Dsui_0158
PS0158c	CCCATCGACTTAATTTATATAATGAAGAAGTAAGGCGACGGCTGG	deletion of Dsui_0158
PS0158d	GNNACTAGTAGGCTCGCTCCCTACTGGCA	deletion of Dsui_0158
pcrSup	GNNGAATTCTCCTCGGAATCCCTAAGCGAACG	complementation of *pcrS*
pcrSdn	GNNGGATCCCCCATTCCGCACAAGGTGTTGC	complementation of *pcrS*
pcrPup	GNNGAATTCCACCCTGCGCTGGTGAGATGA	complementation of *pcrP*
pcrPdn	GNNGGATCCCACCTTGTGCGGAATGGGGT	complementation of *pcrP*
pcrRup	GNNGAATTCGGCAAGGAAGGCCGTATCAGAAA	complementation of *pcrR*
pcrRdn	GNNGGATCCCGGCATGAACCATGACGAGCCT	complementation of *pcrR*
pcrPROMup	GNNGAATTCCTGCGTACTGCTATTGGCGGAA	construction of pRAM62 *pcrA* promoter fragment
pcrPROMdn	GNNGGATCCCTGTTTGACCGAACACCGAAACC	construction of pRAM62 *pcrA* promoter fragment
pcrAup	GNNGGATCCTAGATAAGGAGATGACAATGGTTCAAATG	complementation of *pcrA*
pcrAdn	GNNACTAGTGATCAGCGGATCTGACTAGAAA	complementation of *pcrA*
pcrBup	GNNGGATCCCAGAACAGGATATTAATCATGGCTAATGTG	complementation of *pcrB*
pcrBdn	GNNACTAGTGCCGCTCCATTGCAACATTAGG	complementation of *pcrB*
pcrCup	GNNGGATCCTCTTTCGAGATTGCCCGCCATGAT	complementation of *pcrC*
pcrCdn	GNNACTAGTCCCTGGTGTTACTTTCAGCGTTATTCGG	complementation of *pcrC*
pcrDup	GNNGGATCCGATCCTGTTTAGGAAATGAGGAATCAAATG	complementation of *pcrD*
pcrDdn	GNNACTAGTTCGCGAGTCGGCTCTAAGAATCA	complementation of *pcrD*
cldup	GNNGGATCCCACACAAGGAGATATCCATGACAAAC	complementation of *cld*
clddn	GNNACTAGTTTGAGAGGGACAGCCATTCGGTTA	complementation of *cld*

Chapter 3:

A Conserved Bacterial Mechanism for Modulating the Redox State of Periplasmic Methionine Residues in Response to Reactive Chlorine Species

Abstract

In perchlorate-reducing bacteria, chlorite is generated in the periplasm by perchlorate reductase. The majority of chlorite is converted to chloride and oxygen by chlorite dismutase, but low steady state concentrations of chlorite and hypochlorite exist. The model perchlorate reducer *Azospira suillum* PS contains a unique system for dealing with these reactive chlorine species (RCS). This system consists of a sigma factor/anti-sigma factor system (SigF/NrsF) for specifically sensing the stress, a soluble methionine-rich periplasmic protein (MrpX) that acts as a chloroxyanion scavenger, and a putative periplasmic methionine sulfoxide reductase (YedY1). We have characterized this system by phenotyping an extensive collection of knockout strains, sequencing (RNA-Seq and TN-seq), and biochemical assessment of methionine oxidation (SDS-PAGE and LC-MS/MS). We found that the SigF is specifically activated by RCS and activates transcription of a small regulon centering around *yedY1* and *mrpX*. A *yedY* paralog (*yedY2*) was found using TN-seq to have a similar fitness to *yedY1* despite not being regulated by SigF, and clean deletions confirmed the functional redundancy of these genes. MrpX was found to be strongly induced and rapidly oxidized by RCS. These data suggest a mechanism where MrpX scavenges RCS in the periplasm by sacrificial oxidation, then is regenerated by the methionine sulfoxide reductase activity of YedY. We have also surveyed the phylogenomic distribution of the components of this system, which are conserved in several Proteobacteria of clinical and commercial importance, including uropathogenic *E. coli, Brucella* spp., and root-nodulating members of the Rhizobiales.

Introduction

Bacteria are stressed by reactive chlorine species (RCS) such as hypochlorous acid (HOCl) when encountering the innate immune system and mucosal epithelia of eukaryotes (Ha et al., 2005; Gray, Wholey, and Jakob, 2013). RCS serve a similar antimicrobial function to reactive oxygen species such as hydrogen peroxide and superoxide, but are reactive towards a unique set of biomolecules. One striking example of this is the rate of methionine oxidation by HOCl (3.8×10^7 $M^{-1} \cdot s^{-1}$) (Pattison and Davies, 2001), which is significantly faster than that of hydrogen peroxide (on the order of 10^{-2} $M^{-1} \cdot s^{-1}$) (Pan et al., 2006). Methionine sulfoxide formation thus plays a significant role in the mechanism of bacterial killing by HOCl produced in phagocytes by the myeloperoxidase enzyme (Rosen et al., 2009). Methionine oxidation is also the activating mechanism of the HOCl-specific transcription factor HypT in *E. coli* K12 (Gebendorfer et al., 2012; Drazic et al., 2013). *E. coli* K12 also contains a second transcription factor, NemR, responsive to hypochlorite (Gray, Wholey, Parker, et al., 2013). These transcription factors are conserved beyond *E. coli*, suggesting that hypochlorite defense systems shape bacterial fitness beyond pathogenesis and other interactions with eukaryotes.

Another less well-known biological source of RCS is bacterial dissimilatory chlorate and perchlorate [collectively (per)chlorate] reduction, which is a respiratory process carried out by members of the Proteobacteria known as dissimilatory perchlorate-reducing bacteria (DPRB) (Coates and Achenbach, 2004). These bacteria are obligate respirers that use perchlorate as an electron acceptor, reducing it to chlorite in the periplasm with the enzyme perchlorate reductase (PcrA) (Kengen et al., 1999). Chlorite is then converted to molecular oxygen and chloride by the enzyme chlorite dismutase (Cld), which is conserved in both chlorate and perchlorate reducers (van Ginkel et al., 1996; Ha et al., 2005; Pan et al., 2006; Gray, Wholey, and Jakob, 2013). Chlorite itself is toxic at relatively low concentrations to various organisms including yeast (Kwolek-Mirek et al., 2011), *Desulfovibrio alaskensis* G20 (Carlson et al., 2014), and (per)chlorate-reducing isolates (R. Melnyk and I.C. Clark, unpubished data). However, biochemical studies of Cld suggested that hypochlorite could be an intermediate formed by the enzyme (Lee et al., 2008), and a recent study confirmed production of hypochlorite by the *Nitrospira defluvii* Cld and subsequent oxidative protein damage, which could be ameliorated by addition of hypochlorite trap molecules such as methionine (Hofbauer et al., 2014). This study also measured a rate of hypochlorite release from the enzyme which was in the micromolar range (Hofbauer et al., 2014), suggesting that perchlorate-reducing bacteria could be experiencing RCS stress as a side effect of the metabolism.

The presence of both hypochlorite and chlorite during perchlorate reduction means that RCS stress may be an unavoidable side effect of this metabolism. This has

been previously suggested by us based on the observation in the model DPRB *Azospira suillum* PS (PS) that perchlorate is inhibitory to a Δ*cld* mutant growing on nitrate, but does not inhibit a Δ*pcrA*Δ*cld* mutant's growth on nitrate (Melnyk et al., 2013). We proposed that the epistatic interaction between these two genes was due to accumulation of RCS (specifically chlorite) in the Δ*cld* mutant via inadvertent reduction of perchlorate by PcrA, which was eliminated in the Δ*pcrA*Δ*cld* double mutant. Additionally, oxidative stress defense systems were recently found to be transcriptionally upregulated during chlorate reduction in *Shewanella algae* ACDC, highlighting that RCS may be a ubiquitous problem faced by all chloroxyanion respirers (Clark et al., 2014).

DPRB have a horizontally transferred perchlorate reduction genomic island (PRI) that contains the genes responsible for perchlorate reduction, including *pcrA* and *cld* (Melnyk et al., 2011). Markerless deletions of each of the seventeen genes in the PRI of PS demonstrated that seven have no apparent role in perchlorate reduction yet are conserved in the PRIs of more than one perchlorate reducer (Gray, Wholey, Parker, et al., 2013). Two of these genes encode products similar to the ROS-responsive SigF-NrsF system described in two model Alphaproteobacteria, *Caulobacter crescentus* and *Bradyrhizobium japonicum* (Alvarez-Martinez et al., 2006; Masloboeva et al., 2012), and we hypothesized that SigF-NrsF in PS may also form part of a RCS response during perchlorate reduction.

We found that Δ*sigF* mutants were more sensitive to chlorite stress and had growth defects on chlorate, whereas the Δ*nrsF* phenotypes were reversed. RNA-seq experiments with these mutants identified the SigF regulon as five protein-coding genes, all of which are located in the PRI. These genes included the regulators themselves (*nrsF* and *sigF*), the molybdopterin oxidoreductase components *yedY1* and *yedZ1*, and a gene that encoded a **m**ethionine-**r**ich **p**eriplasmic peptide which we designated MrpX. Using a tagged transposon library, we found that *yedY1* and its paralog located outside of the PRI (*yedY2*) were important during chlorate reduction, and further mutational analysis demonstrated functional redundancy of these two genes. We also used biochemical methods to assess the role that chlorite has in the oxidation state of MrpX and the whole proteome. A phylogenomic survey of YedY reveals that it is associated with pathogenicity islands, arguing for a role for YedY and MrpX in defense against host-derived hypochlorite.

We present these data as evidence that RCS plays a role in (per)chlorate reduction and that some DPRB have a system for ameliorating RCS by modulating the redox status of periplasmic methionine residues. Furthermore, this system is conserved in bacteria with diverse lifestyles, suggesting that RCS may be widely encountered by bacteria in their natural environments.

Results

SigF is required for PS to respond to aerobic chlorite stress

Our recent genetic analysis of *A. suillum* PS demonstrated that 9 of the 17 genes that compose the core of the PRI are dispensable for perchlorate reduction (Melnyk et al., 2013). Of these genes, two are homologous to the *sigF/nrsF* sigma factor/anti-sigma factor system, which has been shown to be responsive to reactive oxygen species in *Caulobacter crescentus* (Alvarez-Martinez et al., 2006; Kohler et al., 2012) and *Bradyrhizobium japonicum*, where it is referred to as *ecfF/osrA* (Masloboeva et al., 2012). In this system, SigF is bound by the membrane protein NrsF until the oxidative signal is sensed, which leads to the release of SigF and transcription of its stress response regulon (Masloboeva et al., 2012; Kohler et al., 2012). This system is an example of the ECF (**e**xtra**c**ytoplasmic **f**unction) sigma factor family, in which an anti-sigma factor transduces environmental signals into transcriptional outputs (Staroń et al., 2009).

The conservation of *sigF/nrsF* in multiple perchlorate reducers was suggestive of a role in defending against reactive intermediates of the metabolism, namely chlorite and possibly hypochlorite. We assayed the ability of Δ*sigF* and Δ*nrsF* mutants to respond to chlorite treatment by adding two aliquots of chlorite during log phase (Figure 1A) and one initial aliquot after inoculation during lag phase (Figure 1B). Addition of chlorite arrested the growth of the Δ*sigF* mutant, despite only slightly inhibiting wild-type PS (Figure 1A). In contrast, the Δ*nrsF* strain challenged with chlorite during lag phase grew identically to unchallenged wild-type PS and better than the challenged wild-type PS (Figure 1B). The Δ*sigF* strain did not have any defect relative to wild-type in the lag phase assay (data not shown), suggesting that these stress conditions represent different fitness regimes.

Although Δ*sigF* or Δ*nrsF* mutations did not have phenotypes during perchlorate reduction in previous work (Melnyk et al., 2013), this may have been due to the use of rich medium with an excess of electron donor. We surmised that oxidative stress may be more pronounced during growth on minimal medium with higher concentrations of electron acceptor, especially chlorate. Chlorate is reduced by PS using the same biochemical pathway that produces the same intermediates, but at a faster rate, which is likely to increase oxidative stress. Using these modified conditions, slight but significant impacts on growth were observed in both the Δ*sigF* and Δ*nrsF* mutants (Figure 2). Like the aerobic chlorite stress phenotypes, Δ*sigF* was deleterious while Δ*nrsF* enhanced growth. The severity of the Δ*sigF* phenotype also seemed to be concentration-dependent, exhibiting a more pronounced defect at higher concentrations of chlorate.

An RNA-seq approach to identify the compact SigF regulon

Because the SigF/NrsF system relies on a single anti-sigma factor that releases SigF in response to chlorite stress, ascertaining the regulon of SigF was possible with two experiments. First, we compared global transcription of chlorite-stressed Δ*sigF* to chlorite-stressed wild-type PS. However, because the Δ*sigF* mutant is defective under these conditions, there may be changes in gene expression unrelated to SigF (e.g. global stress responses not induced in the healthier wild-type strain). To delineate the SigF-specific response, the transcriptomes of unstressed Δ*nrsF* strain and wild-type PS were compared. Only four genes were upregulated in the Δ*nrsF* mutant and downregulated in the Δ*sigF* mutant. These genes compose an operon adjacent to the *sigF-nrsF* operon but on the opposite strand (Figure 3). By our method, *sigF* and *nrsF* were not called as part of this regulon because each was deleted in one experiment, however, SigF does regulate itself and *nrsF* transcriptionally. Many genes were significantly regulated in one experiment but not the other and were thus not deemed part of the SigF regulon. These data likely reflect real transcriptional differences but were ascribed to global physiological changes caused by perturbations in the SigF-NrsF system.

Five PRI encoded genes (Dsui_0154-Dsui_0158) were classified as the functional part of SigF regulon (Figure 3). Although Dsui_0159 was originally included in the regulon, it has a truncated thioredoxin domain and an ATC start codon and is thus likely to be a pseudogene. Dsui_0156 and Dsui_0157 are annotated as *yedY1* and *yedZ1*, which encode a molybdopterin-containing oxidoreductase complex conserved in many members of the Proteobacteria (Loschi et al., 2004). *yedY* homologs are also part of the SigF regulon in *Bradyrhizobium japonicum* and *Caulobacter crescentus*, but their role in the oxidative stress response is not known (Masloboeva et al., 2012; Kohler et al., 2012). Dsui_0158 encodes a small periplasmic protein composed almost entirely of methionines and charged residues, which we have named MrpX (**m**ethionine-**r**ich **p**eptide X). All three genes exhibit statistically significant differential expression in both experiments, but the magnitude varies considerably, with *mrpX* having two orders of magnitude more expression than *yedY1Z1* when counts were normalized by gene length and total number (cf. Figure 4A to 4C,4D). We were somewhat surprised to see that chlorite stress is sufficient to transcriptionally induce many of the genes in the PRI including *pcrA* and *cld* (Figure 5), which encode perchlorate reductase and chlorite dismutase, the two key metabolic enzymes in perchlorate reduction (van Ginkel et al., 1996; Kengen et al., 1999; Melnyk et al., 2013). However, these genes are not under the control of the NrsF-SigF system, as the expression is not altered when either gene is inactivated (Figure 5).

yedY1Z1 and *mrpX* are encoded on the same strand and likely form a single transcriptional unit. Likewise, the *sigF* and *nrsF* genes overlap and form another operon. We focused on the intergenic region between *sigF* and *yedZ* and identified two

locations where the Δ*nrsF* and untreated wild-type transcriptome had increased coverage relative to Δ*sigF* and chlorite-stressed wild-type (Figure 3, inset). Upstream of these two putative transcriptional start sites, we found a conserved promoter similar to the SigF binding motifs reported for both *Caulobacter crescentus* and *Bradyrhizobium japonicum* (Masloboeva et al., 2012; Kohler et al., 2012). The nucleotides conserved in all promoters in all three organisms are indicated in bold (Figure 3). We were unable to find a promoter with this structure anywhere else in the entire PRI, suggesting that these two promoters are the only locations in the PRI where SigF binds.

Although we were only able to identify two promoters, the per-base coverage data display substantial spikes in certain locations, particularly upstream and within the *mrpX* gene (Figure 3). This may be due to actual upregulation, or it could be due to one of the sequence-dependent biases encountered in RNA-seq, namely GC-content or transcript length (Oshlack and Wakefield, 2009; Finotello et al., 2014). We think this is unlikely to be the case, as GC content and transcript abundance have been shown to be positively correlated (Risso et al., 2011), and this region has a lower GC content than the surrounding area. Additionally, the fact that we were only able to identify a single promoter suggests that the *mrpX* region is co-transcribed with upstream genes.

We also used a qPCR assay to quantify upregulation of *mrpX* after chlorite treatment, which also showed a20 to 60-fold increase in transcription with respect to a no treatment control (Figure 4B). This increase was on the same order of magnitude as RNA-seq experiments. Interestingly, *mrpX* was also upregulated to a similar degree by hypochlorite, but not hydrogen peroxide, demonstrating that SigF is not activated indiscriminately by all reactive oxygen species (Figure 4B).

***mrpX* and *yedY* are each essential for chlorite resistance in the Δnrs*F* knockout, but are not required for optimal growth on chlorate**

The gene expression data suggested that the resistance of the Δ*nrsF* mutant is due to the expression of only three protein-coding genes: *yedY1Z1* and *mrpX*. To see if these genes were necessary for chlorite resistance, we created Δ*nrsF*Δ*yedY1* and Δ*nrsF*Δ*mrpX* strains and grew them under chlorite stress. Both strains were as sensitive as wild-type to chlorite, demonstrating that the transcriptional upregulation of each gene is required for Δ*nrsF* chlorite resistance (Figure 6A). However, when these strains were grown on chlorate, they still displayed the improved growth seen in the Δ*nrsF* mutant (Figure 6B), suggesting possible redundancy within the regulon under this condition.

BarSeq fitness profiling indicates that the SigF regulon is important during chlorate reduction

Our previous genetic analysis showed that PRI genes are not important for aerobic growth or denitrification, but that eight are essential for perchlorate reduction (Melnyk et al., 2013; Hofbauer et al., 2014). To search for genes with more subtle phenotypes, including those outside the PRI, a saturated transposon mutant library containing unique TagModule barcodes in each insertion was generated (Oh et al., 2010; Melnyk et al., 2013). Barcodes were mapped to their respective gene, and the fitness of each strain was determined by comparing barcode abundance before and after specific treatments. Gene fitness values represent the fitness benefit (positive score) or fitness defect (negative score) associated with inactivation of the gene under a given condition (Melnyk et al., 2011; Deutschbauer et al., 2011). This approach is called BarSeq, and is being used to rapidly generate gene fitness data for diverse microorganisms. We used the mutant pool generated in PS to profile gene fitness in response to oxidative stress and on different electron acceptors as well as under different growth conditions. We used minimal media with 30 mM lactate to test the different electron acceptors (oxygen, nitrate, chlorate, perchlorate) and rich ALP media to test the stress conditions (chlorite, hydrogen peroxide, and an unstressed control).

BarSeq data recapitulated previous clean deletion phenotypes of genes in the PRI; eight of the 17 genes (*pcrABCDPSR* and *cld*) are essential for perchlorate reduction and this is reflected by the significant fitness defect for all eight genes under the perchlorate and chlorate conditions (Figure 7A). The fitness defects for *sigF*, *yedY*, *yedZ*, and *mrpX* were milder than the essential genes and seemed to be chlorate-specific, as there was no measurable defect during perchlorate reduction (Figure 7A, 7B). This is consistent with the previous observation that deletions of these genes were not defective when grown on rich media with perchlorate as the electron acceptor (Melnyk et al., 2013). The fitness defect for *sigF* was significantly greater than the defect associated with any individual gene in the regulon, supporting the hypothesis that there is some degree of redundancy from a fitness perspective within the *sigF* regulon (Figure 7B).

Because the *sigF* mutant only has a mild defect on chlorate despite the basal transcription of the SigF regulon, we hypothesized that there may be redundant mechanisms of chlorite resistance outside the SigF regulon that are allowing growth. To find these genes, we used the **H**ierarchical **Cl**ustering (HCL) algorithm within the MeV software package to find genes with fitness profiles that clustered with the chlorate-specific defect seen in the members of the SigF regulon (Eisen et al., 1998; Saeed et al., 2003). We extracted a cluster of eleven genes from the analysis that contained all three genes in the SigF regulon (Figure 7C). Two of the other genes in this cluster were homologous to *yedY1* and *yedZ1* (Dsui_1300 and Dsui_1299, hereafter, *yedY2* and *yedZ2*), providing a possible target for double mutant studies to test for synthetic lethality. We thus created a clean deletion of Δ*yedY2* as well as a double

knockout of Δ*sigF*Δ*yedY2*. The Δ*yedY2* single mutant had no obvious phenotype when grown on chlorate, but the Δ*sigF*Δ*yedY2* mutant had a very strong defect on chlorate (Figure 8). Like the Δ*sigF* mutant, the severity of the defect was proportional to concentration, with the double knockout exhibiting no growth at 20 mM chlorate concentration.

We also stressed the pool with an initial chlorite spike in order to find genes important for chlorite resistance. However, genes in the *sigF* regulon and *sigF* itself were not defective in this assay and neither was *cld*, which is a known determinant of chlorite resistance both in native perchlorate reducers and engineered strains of *Caulobacter crescentus* (Charlie Huang and Kathleen Ryan, unpublished data). One possible explanation for this discrepancy is the lack of an induction period, as the stress in the BarSeq experiments was restricted to single initial spike, rather than exposure over a long time period. If SigF was not given enough time to induce its regulon, inactivation of *sigF* or its regulon would not have an effect on fitness. This concept may also explain why a Δ*sigF* mutant is no different from wild-type when treated during lag phase with a single chlorite spike (Figure 1B), but is sensitive when treated with multiple spikes mimicking chronic chlorite exposure (Figure 1A).

Although SigF regulon mutants were not defective under the chlorite stress, there were some genes that did have defects specific to the chlorite stress condition (Figure 1C). These ten genes were clustered together in the HCL analysis and included one gene from the PRI (Dsui_0153). This gene encodes a cupin domain protein that is not essential for (per)chlorate reduction. Mutations in a subset of genes (Dsui_1304, Dsui_1652, Dsui_1653) in this cluster were sensitive to chlorite stress, but had fitness benefits when grown on perchlorate and chlorate. Dsui_1304 encodes a *c*-type cytochrome, and Dsui_1652-Dsui_1653 encode FeoAB, a ferrous iron transporter (Cartron et al., 2006). Control of cytoplasmic iron is one of the major outputs of one of the hypochlorite-sensitive transcription factor HypT, and may represent a conserved strategy for dealing with RCS (Drazic et al., 2013).

The methionine residues of MrpX are vulnerable to oxidation by chlorite

We hypothesized that the physiological function of MrpX is to provide a source of methionine residues that can be sacrificially oxidized to remove RCS from the periplasm. To monitor the oxidation state of MrpX during chlorite treatment, we took advantage of the fact that methionine oxidation results in a disproportionate shift in protein igration in denaturing polyacrylamide gel electrophoresis (Le et al., 2008). We used allelic replacement to add a C-terminal myc-tag to the chromosomal *mrpX* in order to preserve native transcriptional regulation. This replacement was also performed in the Δ*nrsF* background to generate a positive control strain that constitutively overexpressed a tagged version of MrpX. We grew cultures of these two strains, as well

as of an untagged wild-type strain and used a simple chloroform extraction to get periplasmic protein fractions from all three strains (Ames et al., 1984; Alvarez-Martinez et al., 2006; Kohler et al., 2012). We were able to detect MrpX:Myc via Western blotting against the Myc tag and we could also see the MrpX:Myc protein in the Δ*nrsF* background via a simple total protein stain (Figure 9A).

We grew wild-type PS containing the tagged *mrpX* gene and then spiked in chlorite, harvesting samples for periplasmic protein extraction every ten minutes. When these samples were analyzed by Western blotting, we observed a gradual accumulation of MrpX initially in the oxidized form which shifted downwards over time, reaching the same level as the reduced protein control after 50-60 minutes (Figure 9B). To show that this effect is indeed caused by chlorite-dependent oxidation, we reversed the shift by extracting periplasmic protein from the Δ*nrsF mrpX*::*myc* strain and treating with increasing amounts of chlorite *in vitro* (data not shown).

Chlorite increases methionine oxidation across the proteome

Hypochlorite is a well-known oxidizer of methionine residues both *in vivo* and *in vitro* (Hawkins et al., 2003; Staroń et al., 2009; Rosen et al., 2009; Masloboeva et al., 2012; Kohler et al., 2012), but we were not aware of chlorite being definitively shown to be a potent oxidizer of methionine residues. We used aerobically grown wild-type and Δ*sigF* cells to look at the effect of chlorite treatment on the global methionine redox state. In both genotypes, 200 μM chlorite is sufficient to substantially elevate the percentage of methionine oxidized (Figure 10). However, there appears to be no significant difference in global methionine oxidation between wild-type and Δ*sigF* under either of the conditions tested (Figure 10).

Exogenous methionine improves growth on perchlorate and chlorate

The Δ*nrsF* mutant overexpresses the *sigF* regulon, which includes *mrpX*, and exhibits improved growth on chlorate. This suggests that the exogenous addition of methionine may also provide a growth benefit to wild-type cells. We added varying amounts of methionine to minimal media with lactate and grew wild-type cells using chlorate or perchlorate as the electron acceptor. Methionine was inhibitory above a certain threshold (~1 mM), but at lower concentrations shortened the lag time associated with growth under perchlorate-reducing conditions (Figure 11A). A more modest but significant improvement was observed for growth with chlorate (Figure 11C).

Biochemical studies involving the YedY protein from *E. coli* suggest that it has methionine sulfoxide reduction activity (Loschi et al., 2004), and we reasoned that a Δ*yedY1*Δ*yedY2* mutant would not benefit from exogenous methionine if redox cycling of the free amino acid was involved. Consistent with this hypothesis, the double mutant

had identical growth profiles to untreated wild-type after addition of 200 μM methionine on either chlorate or perchlorate (Figure 11B, 11D).

YedYZ and MrpX are broadly conserved in diverse bacteria

YedY is broadly distributed throughout the Proteobacteria and forms two distinct clades that are part of a larger sulfite oxidase fold family (Workun et al., 2008). We phylogenetically classified the YedY paralogs from PS and observed that they fell into separate clades. One clade contained the YedY2 sequence from the PS chromosome, as well as the YedY proteins from *E. coli* and *C. crescentus*. The other clade contained the YedY from the PS PRI. Our initial exploration of the sequences from the two clades suggested that the YedY2 clade was distributed across the Proteobacteria and perhaps has a possible 'housekeeping' role, whereas the clade containing YedY from the PRI was restricted to certain subgroups within the Proteobacteria and was frequently found associated with markers of horizontal gene transfer (HGT). The phylogeny of the housekeeping clade also mimicked the major order taxonomic divisions within the Proteobacteria (data not shown), while the HGT clade featured interspersed taxa from several Proteobacterial orders. One obvious difference that demarcates the two clades is the organization of the *yedYZ* operon: in the housekeeping clade *yedY* is upstream of *yedZ*, while in the HGT clade the order is reversed.

We generated a phylogeny by concatenating the HGT clade YedY sequences with the YedZ sequences from the adjacent genes to improve the tree quality, and rooted the tree using the housekeeping sequences of YedYZ from *Caulobacter crescentus*, *Azospira suillum* PS, and *E. coli* CFT073 (Figure 12). We then inspected the regions surrounding the *yedY* genes from this clade to identify common linked genes, which are displayed for each taxon in Figure 12. *mrpX* was found near *yedY* 14 times in this clade and *sigF*/*nrsF* 6 times. Three *yedYZ* sequences were next to genes for a histidine kinase system, which may be a regulatory system parallel to *sigF*/*nrsF*. 5 *yedY* sequences were nearby putative thioredoxin genes similar to Dsui_0159, the putative thioredoxin in PS. Finally, in three instances we found *yedY* linked to the methionine sulfoxide reductases *msrA* or *msrB*.

While the housekeeping *yedY* is distributed throughout the Proteobacteria (Figure 13A), the HGT *yedY* is enriched in a few subsets of bacteria of broader importance (Figure 13B). For instance, only six *E. coli* strains of the 43 included in our dataset have *yedY* homologs from the HGT clade in their genomes (represented by the Q8FE04 taxon in Figure 11), while all have the housekeeping *yedY* (Figure 13B). Five of these strains are either uropathogenic *E. coli* (UPEC) or agents of asymptomatic bacteriuria (strain 83972 (Hull et al., 1999)) and are the only such isolates in our dataset. In these five strains, the *yedZYmrpX* operon is near an operon of hemolysin genes and is thus part of the pathogenicity island PAI-II in strain 536 (Dobrindt et al., 2002) and PAI-

I in CFT073 (Guyer et al., 1998; Lloyd et al., 2007). In the sixth *E. coli* isolate H10407, which is an enterotoxigenic (ETEC) strain, the *yedYZmrpX* operon is located on a plasmid. This strain is a unique case among pathogenic *E. coli*, as it has no pathogenicity islands on its chromosome, but rather has several plasmids which confer virulence traits (Crossman et al., 2010).

We also observed conservation of both *yedY* types in the Brucellaceae family (Figure 13B). All 12 of the *Brucella* strains had both a housekeeping *yedY* and an HGT *yedY*, whereas the soil bacterium *Ochrobactrum anthropi* had only the housekeeping *yedY*. *O. anthropi* and *Brucella* spp. shared a recent common ancestor before the *Brucella* ancestor transitioned to the pathogenic lifestyle via massive genome reduction and acquisition of virulence determinants (Wattam et al., 2013). This means that the housekeeping *yedY* was conserved despite 30% of the ancestral genome being lost in *Brucella* relative to *O. anthropi*. Additionally, the *Brucella* lineage acquired an *yedYZmrpX* operon which is now located next to an enterobactin siderophore operon that has been identified as a virulence determinant unique to *Brucella* (Bellaire et al., 2003; Wattam et al., 2013).

Finally, we surveyed the genes predicted to be regulated by SigF in *Bradyrhizobium japonicum*, which includes a member of the HGT *yedY* clade and a gene predicted to encode a methionine-rich peptide (Masloboeva et al., 2012). Unlike the genomic organization in UPEC and PS, the genes are scattered throughout the *B. japonicum* genome (Figure 14). The *B. japonicum* SigF regulon is more extensive than the regulon in PS, including several methionine sulfoxide reductases from the MsrA and MsrB families (Masloboeva et al., 2012). We displayed one such operon here, which includes a *msrA* gene with a predicted signal sequence which is relatively uncommon (Figure 14).

Discussion

This characterization of the SigF-NrsF system establishes the role of RCS during perchlorate reduction and identifies two proteins, YedY and MrpX, as being determinants of oxidative stress resistance. This system was characterized by using classic genetics techniques to generate mutants with phenotypes under chlorite stress. We then used RNA-seq and BarSeq approaches to identify and characterize the stress response regulon, which centers around methionine redox chemistry in the periplasm. We also used biochemical methods to show that chlorite stress results in oxidation of methionine residues both in MrpX and across the whole proteome. Also, we showed that addition of exogenous methionine *in vivo* improves growth on (per)chlorate in a YedY-dependent manner. Finally, we present the phylogenomic evidence for YedY

being both a common housekeeping pathway in the Proteobacteria and perhaps a fitness determinant for bacteria with host-associated lifestyles.

Our RNA-seq analysis definitively shows that the output of SigF in PS is limited, as it activates transcription of only two promoters controlling five protein-coding genes. This is in contrast to other ECF factors such as RpoE in *E. coli*, which controls a regulon consisting of at least 20 promoters throughout the genome (Dartigalongue et al., 2001). However, the genetic linkage of the *sigFnrsF* operon with its regulon (*yedY1,* yedZ2, and *mrpX*) and their location within the horizontally transferred PRI offers an explanation. For a mechanism of transcriptional regulation to be an adaptive part of a horizontally transferred genomic island, there must either be promoters that it controls already in the recipient genome, or it must control promoters within the island. In the case of SigF in PS, it only controls PRI genes. This is in contrast to the previously characterized SigF regulons, which control promoters scattered throughout the genome (Masloboeva et al., 2012; Kohler et al., 2012). Perhaps given enough evolutionary time, SigF-dependent promoters elsewhere in the PS genome could emerge. Our BarSeq data shows that there are many genes in the PRI which are specifically important for either acute chlorite stress or growth on (per)chlorate but are not under the control of SigF.

One of these genes, Dsui_0153, lies within the PRI and has a cryptic role. Its transcriptional profile is similar to *pcrA* and *cld* in that it is activated by chlorite stress but not controlled by SigF. Interestingly, inactivation of this gene, which encodes a cupin domain protein (Dunwell et al., 2004), leads to a defect under acute chlorite stress, but not during growth on chlorate or perchlorate. This is in contrast to the fitness profile of *sigF*, which is important on chlorate but not under acute chlorite stress during lag phase. This discrepancy is important to consider as the oxidative stress community moves towards whole-genome fitness analysis as we carried out here. The reactive nature of oxidative species means that a single initial spike of stressor may not have a very long half-life, and therefore fitness profiling will reflect sensitivity or resistance of the initial population to killing rather than a true fitness defect or benefit over many generations. This becomes important when considering the phenotype of genes like *sigF* which control transcription of fitness determinants. Therefore, we recommend that experiments with an initial spike of a stressor be coupled with repeated spikes or chronic exposure at a much lower dose for more biologically relevant results.

The concept of the methionine residues of a protein acting as antioxidants was first proposed by Levine *et al.*, who showed that glutamine synthase from *E. coli* had several methionine residues which were surface exposed and could be oxidized with little effect on enzyme function (Levine et al., 1996). Thus, the function of methionine sulfoxide reductases is two-fold: not only do they repair non-functional proteins, but they also regenerate exposed methionines which act as sinks for oxidative stress (Ezraty et al., 2005). We propose that MrpX is primarily an oxidative stress sink; its lack of

sequence conservation and high proportion of charged residues mean that it likely is natively unfolded, which may allow preferential oxidation and reduction of its methionine residues (Tarrago et al., 2012). A simple search of a large sequence database suggested that there are many small peptides in both prokaryotes and eukaryotes that contain many methionines and charged residues (R. Melnyk, unpublished data). One of these is the hydrophilin protein Sip18 from *Saccharomyces cerevisiae*, which has a role in desiccation and oxidative stress tolerance (Rodríguez-Porrata et al., 2012; López-Martínez et al., 2012).

The function of the YedY proteins was not initially clear, given that the only described member of this family does not have a known biological function. However, the excellent biochemical work done on this protein from *E. coli* showed that it is likely a reductase, not an oxidase, and is able to reduce free methionine sulfoxide *in vitro* (Loschi et al., 2004; Brokx et al., 2005; Havelius et al., 2011). We propose that this is its role *in vivo* as well based on its contribution to fitness in the presence of RCS and the fact that the double Δ*yedY1*Δ*yedY2* mutant does not benefit from additional methionine. We propose that this role extends beyond perchlorate reducers based on observations in other bacteria. First, a microarray study on *E. coli* O157:H7 shows upregulation of *yedYZ* in response to hypochlorite, but not hydrogen peroxide (Wang et al., 2009). Secondly, fitness profiling of two *Shewanella* strains (MR-1 and SB2B), shows that *yedYZ* are dispensable for hundreds of conditions, with the lone exception of chlorite stress where insertions lead to a strong growth defect (M. Price, unpublished data). A recent study from our lab showed transcriptional upregulation in the chlorate-reducing strain *Shewanella algae* ACDC as well (Clark et al., 2014). The genomes of these organisms contain only a housekeeping YedY paralog, thus we expect that the function is conserved between both clades from our phylogenomic analysis.

Future biochemical studies should bear out the precise function of YedY and answer crucial questions about its activity such as stereoselectivity and preference for free methionine sulfoxide versus peptide methionine sulfoxide (Ezraty et al., 2005). Additionally, the mechanism of toxicity of chlorite needs to be explored in more biochemical detail. We were unable to identify global shifts in methionine oxidation in our mutants, but the mechanism of protection may be specific. For example, MrpX may interact closely with Cld, the point source of hypochlorite production in DPRB. This would prevent the oxidation of key residues in the Cld and subsequent protein inactivation, thereby enhancing growth in the same way that free methionine enhances activity of the pure protein (Hofbauer et al., 2014). Such a change would not radically shift the global methionine oxidation state of the cell, and thus would be undetected in the global proteomics experiments carried out in this study.

The conservation of the *yedYZmrpX* operon in UPEC and *Brucella* spp. suggests that methionine redox cycling plays a role in pathogenesis, perhaps in ameliorating

hypochlorite stress caused by myeloperoxidase. Virulence is greatly attenuated but growth is unimpaired in UPEC strain 536 when the PAI-II island containing *yedYZmrpX* is deleted (Brzuszkiewicz et al., 2006). Many other genes with known importance to UPEC such as hemolysins and fimbriae are in this island, so targeted deletions of *yedYZmrpX* are necessary. Redundant biochemical mechanisms may also play a role as they do in PS, where deletion of a paralogous *yedY* gene was needed before substantial phenotypes were observed. Additionally, the histidine kinase and response regulator genes upstream of *yedZ* in UPEC are syntenically conserved in various strains with the HGT *yedY* (Figure 14). This histidine kinase/response-regulator could play a role much like SigF and NrsF: that is, providing environmental sensing and transcriptional regulation to a horizontally transferred operon. This regulatory mechanism would be important to UPEC, as the uropathogenic life cycle involves various tissue environments, only some of which would be expected to harbor phagocytes and thus reactive chlorine species.

The signatures of horizontal gene transfer in several of the genomes we analyzed were unmistakable. The *yedYZmrpX* operon lies within the 25-kb region defined as the PRI in *Azospira suillum* PS, and it is within the island PAI-II in UPEC strain 536 that has been documented to undergo conjugal transfer under laboratory conditions (Schneider et al., 2011). However, the genome-wide distribution of the *sigF* gene and the SigF regulon in *Bradyrhizobium japonicum* is more suggestive of vertical descent, as none of these features are part of the symbiosis genomic island in *B. japonicum* (Kaneko et al., 2002; Masloboeva et al., 2012). Additionally, the genomic region surrounding the *sigFnrsF* and *yedYZmrpX* operons in *Brucella* also does not appear to be part of a genomic island. The prevalence of the HGT *yedY* in the Rhizobiales suggests that it may have arisen there, or have been transferred long ago, such that fixation and partitioning of the genes could occur. This also implies that the function of the YedY-MrpX system is of importance to certain Rhizobiales. This is perhaps unsurprising, as *Brucella* spp. are facultative intracellular pathogens that likely encounter reactive chlorine species at least periodically, despite their ability to attenuate the initial innate immune response (Roop et al., 2009).

Unlike animals, plants are not known to make hypochlorite in response to rhizobial interaction, where the oxidative species hydrogen peroxide and superoxide are dominant (Santos et al., 2001). It is possible that certain plants have enzymes with haloperoxidase activity that are currently undiscovered. Alternatively, hypochlorite stress could be self-inflicted, that is, produced by the invading bacteria when faced with the amounts of hydrogen peroxide produced by plant immunity. A secreted chloroperoxidase is induced by oxidative stress in *Sinorhizobium meliloti*, although its physiological function is unknown (Barloy-Hubler et al., 2004). The presence of robust systems for methionine redox homeostasis in the periplasm may mean that converting

hydrogen peroxide to hypochlorite is an adaptive mechanism for plant-associated rhizobia, perhaps due to its inhibition of nearby microbes that lack the YedY-MrpX system.

In this paper, we set out to characterize SigF as a possible regulator of oxidative stress during perchlorate reduction. The mechanism by which this happens depends on upregulation of a methionine-rich protein and a putative methionine sulfoxide reductase, which are conserved in many other bacteria. This conservation suggests that bacteria with lifestyles beyond enteric commensals and pathogens require the ability to specifically respond to RCS. We present perchlorate-reducing bacteria as one such group, but we suspect future work will uncover others.

Materials and Methods

Culturing methods for growth curves and other experiments

All strains of *Azospira suillum* PS were revived from freezer stocks by streaking for single colonies on agar plates containing ALP medium, a rich media created for routine aerobic culturing of freshwater perchlorate reducers (Melnyk et al., 2013). ALP was also the medium used for several growth curves, but in other cases a minimal media was used. This minimal media is composed of the same ingredients as ALP, with the exception of all electron donors (lactate, acetate, and pyruvate) and yeast extract. When required for certain strains, kanamycin was added to media at a final concentraion of 50 μg/mL. All anaerobic growth curves were performed with four biological replicates in 96-well plates in a volume of 300 μL with an overlay of 75 μL mineral oil. All strains were grown overnight then diluted back to an OD_{600} of 0.9-1.1 for inoculation. The initial inoculum was a 30-fold dilution into the new condition for a starting optical density of ~0.03. All plates were prepared aerobically on the benchtop, then brought into an anaerobic glove bag (Coy Lab Products, Grass Lake, MI) containing nitrogen and 2-3% hydrogen, where they were loaded into a Sunrise plate reader (Tecan Group Ltd., Männedorf, Switzerland) and incubated without shaking at 37°C for 48-72 hours.

For certain experiments involving chlorite stress, cells were grown aerobically using several methods. For aerobic chlorite stress growth curves two methods were used. One method involved spiking in two successive aliquots of chlorite during log phase (Figure 1A). This was done in 30-mL aerobic glass tubes containing 10 mL of media which were shaken at 250 rpm. The second method was done in 96-well plates with chlorite added immediately after cells were diluted back in the new condition to an OD of ~0.03 (Figures 1B, 5A). Each well contained a volume of 150 μL, and the plates were sealed using a translucent adhesive plate seal, and incubated with shaking in a Tecan Sunrise on the benchtop set to 37°C. For the RNA-seq experiments, we

attempted to minimize the ROS created during aerobic respiration by culturing the cells at a lower oxygen tension. This was done by filling a 50-mL polypropylene Falcon tube (BD, Franklin Lakes, NJ, USA) with 30 mL of ALP media and screwing the cap on partway before shaking at 250 rpm at 37°C. Optical density was measured periodically and a single spike of 200 μM chlorite was added to three wild-type replicates and three Δ*sigF* replicates when the OD_{600} reached 0.3-0.4. The cells were returned to the incubator for 30 minutes to allow for induction before RNA was harvested. Cells for the MrpX:myc western blotting, proteomics, and qPCR experiments were grown under the same conditions as the RNA-seq samples. For the qPCR experiments, 100 μM chlorite, 100 μM hypochlorite, and 1 mM hydrogen peroxide were used as the stressors. Chlorite stocks were prepared by dissolving solid sodium chlorite in water. Hypochlorite and hydrogen peroxide stocks were made by diluting concentrated solutions of sodium hypochlorite and hydrogen peroxide. All ROS/RCS stocks were prepared fresh for each experiment.

Strain construction and complementation

Gene deletions in *Azospira suillum* PS were created as previously described (Melnyk et al., 2013). Briefly, a suicide vector carrying a kanamycin marker was electroporated into aliquots of electrocompetent cells before recovery and selection on ALP-Kan media. Positive recipients were grown overnight on unselective media, then plated on ALP media containing 6% sucrose to counterselect against the presence of the *sacB* gene on the integrated plasmid. Sucrose-resistant colonies were then screened using primers flanking the gene of interest before freezer stocks were made. Double mutants were performed similarly, except starting with electrocompetent aliquots of a single mutant strain rather than wild-type PS. Suicide vectors were constructed as previously described using conventional cloning techniques (Melnyk et al., 2013).

To tag the *mrpX* gene with the myc tag, a 'round-the-horn site-directed mutagenesis strategy was used (Sean Moore, openwetware.org). Briefly, a region containing *mrpX* was amplified from PS gDNA using the flanking primers used for making the *mrpX* deletion construct (PS0158a and PS0158d). This was cloned into a vector using the pCR-Blunt II-TOPO kit (Life Technologies). Next, primers were designed that would bind at the C-terminus of the *mrpX* gene but oriented in different directions. One of these primers contained the sequence for the myc-tag. These primers were used to generate a linear PCR product using 10 ng of the *mrpX* TOPO vector as a template. This linear PCR product was circularized using T4 DNA ligase using an overnight cycle that gradually decreased the reaction temperature from 22°C to 4°C. The circularized plasmid was then transformed into chemically competent *E. coli*. Positive transformants were grown up and plasmid DNA was harvested for sequencing. The myc-tagged *mrpX* insert was then subcloned into the pNPTS138

suicide vector which was then introduced into both Δ*mrpX* and Δ*mrpX*Δ*nrsF* strains using the sucrose counterselection allelic exchange method. A complete list of strains, primers, and plasmids used in this study can be found in Tables 1-3.

Determining the SigF regulon using RNA-seq

The cells for the RNA (described under "Culturing methods") were removed from the incubator and placed on ice prior to centrifugation at 4000 rcf at 4°C for 15 min. The supernatant was removed and the cells were resuspended in 1.5 mL TRIzol (Life Technologies). RNA was then extracted using the manufacturer's method. DNA was removed using the TURBO DNA-free Kit (Life Technologies) according to manufacturer's method. RNA quantity and purity was assessed using a NanoDrop ND-2000 (Thermo Scientific). Illumina library prep and sequencing was carried out at the Vincent J. Coates Genomics Sequencing Laboratory. Ribosomal RNA was removed using the Ribo-Zero rRNA removal kit (Epicentre) prior to cDNA synthesis and library construction using the Apollo 324 (WaferGen Biosystems). Shearing of DNA and library quality checks were performed using a Covaris S2 and an Agilent 2100 Bioanalyzer. Samples were multiplexed on a single Illumina HiSeq 2000 lane for single-read sequencing of 50-bp reads.

For calling differential expression of genes, a file containing the sequences of all PS genes was used as the template for Bowtie2 (Langmead and Salzberg, 2012) and the output was streamed to eXpress (Roberts and Pachter, 2013) for quantification of reads mapped to individual genes. Raw counts per gene from eXpress were then used as input to the R package DESeq for statistical inference of differential expression between the two pairs of experimental conditions (Anders and Huber, 2010). We considered any gene with an adjusted p-value < 0.05 between conditions to have significant differential expression. The regulon was defined as any gene that was positively regulated in the Δ*nrsF* mutant relative to untreated wild-type, and also negatively regulated in the Δ*sigF* mutant relative to the chlorite-treated wild-type. In order to create the per-base coverage data for the PRI, we again used Bowtie2 to generate an alignment file, but this time we used the entire PS genome. The samtools suite was then used to sort and generate the per-base coverage data using pileup (H. Li et al., 2009). To visualize transcription, RNAseq data was mapped to the entire PS genome with Bowtie2, per-base coverage determined with mpileup (part of the SAMtools suite (H. Li et al., 2009)), normalized, and the average of replicate treatments was plotted over genomic features to generate the core of Figure 3.

qPCR experiments

RNA was extracted as it was for the RNA-seq sample prep. After quantification, 5 μL of total RNA (~250 ng) was used as the template for a SuperScript II Reverse

Transcriptase reaction (Life Technologies) according to manufacturer's directions using random hexamer primers. *rpoB* was chosen as the housekeeping control gene based on its observed stable expression under all of the RNA-seq conditions tested. Primers for *rpoB* and *mrpX* were designed and their primer efficiency was calculated using a genomic DNA dilution series using the Maxima SYBR Green 2x Master Mix (Thermo Scientific) on the StepOnePlus (Life Technologies). The actual qPCR experiment was performed with both biological and technical triplicates and 1 uL cDNA as template for all reactions (both *rpoB* and *mrpX*). Relative quantification was performed using the Pfaffl method with the specific primer efficiencies for each primer pair incorporated for normalization (Pfaffl, 2001). All calculations were performed using the StepOnePlus software.

SDS-PAGE mobility shift assay

The periplasmic protein extraction was performed based on a previously published method using chloroform (Ames et al., 1984). Ten mL of PS cells were centrifuged at 4000 rpm for 15 minutes then the supernatant was discarded and the cells were resuspended to a volume of ~0.3 mL. 30 μL of chloroform was added and then cells were incubated with gentle shaking at room temperature for 15 minutes followed by 10 minutes of centrifugation at 6000 rcf, also at room temperature. The resulting aqueous phase was then transferred to a clean tube and 4x NuPage LDS Sample Buffer (Life Technologies) was added before boiling the samples at 95°C for 5 minutes. For the *in vitro* assay, various amounts of chlorite were added to the aqueous protein before addition of the sample buffer. 10 μL of the samples were loaded on two 10% NuPage Bis-Tris gels (Life Technologies) and subjected to electrophoresis in a MES buffer for 35 minutes at 200 volts. One gel was stained for total protein using SimplyBlue SafeStain while the other gel was transferred for one hour at 4°C and 200 mA to an Immobilon-PSQ PVDF membrane (Millipore). The membrane was then blocked for 1 hour in 5% nonfat dry milk in 0.5% Tween in Tris-buffered saline (Tween-TBS). The primary antibody (Anti-Myc Tag, clone 4A6, Millipore) was then added at a 1:1000 ratio and incubated for 1 hour at room temperature. The membrane was then washed three times in Tween-TBS. The secondary antibody (Goat Anti-Mouse IgG, HRP conjugate, Millipore) was then added in 5% milk/TBS-Tween and also incubated at room temperature for 1 hour, followed by three more TBS-Tween washings. The Immobilon Western Chemiluminescent HRP Substrate (Millipore) was then added to the membrane, which was imaged over 3 minutes using a Chemidoc (Bio-Rad).

Global methionine oxidation mass spectrometry

After centrifugation, cells were resuspended in 100 mM ammonium bicarbonate and 2 mM methionine to quench any remaining chlorite. Cells were then sonicated for

three cycles of 30 sec each at 4°C, followed by addition of 500 ng trypsin (Trypsin Gold, Promega) and overnight incubation at 37°C. Digested samples were centrifuged at 20000 rpm for 20 minutes, acidified to a final concentration of 0.1% trifluoroacetic acid (TFA), and concentrated with OMIX Tips (Agilent Technologies). Samples eluted with 85% acetonitrile/0.1% TFA were vacuum centrifuged to a volume 30 μl and peptides were identified using liquid chromatography tandem mass spectrometry. Trypsin-digested proteins were analyzed using a Dionex UltiMate3000 RSLCnano liquid chromatograph that was connected in-line with an LTQ Orbitrap XL mass spectrometer equipped with a nanoelectrospray ionization source (nanoESI; Thermo Fisher Scientific, Waltham, MA). The LC was equipped with a C18 analytical column (Acclaim® PepMap RSLC, 150 mm length × 0.075 mm inner diameter, 2 μm particles, 100 Å pores, Thermo) and a 1 μL sample loop. Acetonitrile (Fisher Optima grade, 99.9%), formic acid (Pierce, 1 mL ampules, 99+%), and water purified to a resistivity of 18.2 MΩ•cm (at 25 °C) using a Milli-Q Gradient ultrapure water purification system (Millipore, Billerica, MA) were used to prepare mobile phase solvents for liquid chromatography-tandem mass spectrometry. Solvent A was 99.9% water/0.1% formic acid and solvent B was 99.9% acetonitrile/0.1% formic acid (v/v). Samples contained in polypropylene autosampler vials with septa caps (Wheaton, Millville, NJ) were loaded into the autosampler compartment prior to analysis. The autosampler compartment was maintained at 4 °C. The elution program consisted of isocratic flow at 5% B for 4 min, a linear gradient to 30% B over 98 min, isocratic flow at 95% B for 6 min, and isocratic flow at 5% B for 12 min, at a flow rate of 300 nL/min. The column exit was connected to the nanoESI emitter in the nanoESI source of the mass spectrometer using polyimide-coated, fused-silica tubing (20 μm inner diameter × 280 μm outer diameter, Thermo). Full-scan mass spectra were acquired in the positive ion mode over the range m/z = 350 to 1600 using the Orbitrap mass analyzer, in profile format, with a mass resolution setting of 60,000 (at m/z = 400, measured at full width at half-maximum peak height). In the data-dependent mode, the eight most intense ions exceeding an intensity threshold of 30,000 counts were selected from each full-scan mass spectrum for tandem mass spectrometry (MS/MS) analysis using collision-induced dissociation (CID). MS/MS spectra were acquired using the linear ion trap, in centroid format, with the following parameters: isolation width 3 m/z units, normalized collision energy 28%, default charge state 2+, activation Q 0.25, and activation time 30 ms. Real-time charge state screening was enabled to exclude singly charged ions and unassigned charge states from MS/MS analysis. To avoid the occurrence of redundant MS/MS measurements, real-time dynamic exclusion was enabled to preclude re-selection of previously analyzed precursor ions, with the following parameters: repeat count 2, repeat duration 10 s, exclusion list size 500, exclusion duration 90 s, and exclusion mass width 20 ppm. Data acquisition was controlled using Xcalibur software (version 2.0.7

SP1, Thermo). Raw data files were searched against the *Azospira suillum* PS protein database using Proteome Discoverer software (version 1.3, SEQUEST algorithm, Thermo) for tryptic peptides (i.e., peptides resulting from cleavage at the C-terminal end of lysine and arginine residues) with up to three missed cleavages and methionine sulfoxide as a variable post-translational modification. A decoy database was used to characterize the false discovery rate and the target false discovery rate was 0.01 (i.e., 1%).

***Phylogenomic analysis of* YedY**

The YedY sequence from PS (Dsui_0157) was used as the query for a phmmer (Eddy, 1998) search against complete proteomes from a local database downloaded from HAMAP in August 2013 (Lima et al., 2009). We then used CD-HIT to cluster highly similar protein sequences to reduce the size of the dataset (W. Li and Godzik, 2006). An alignment was generated using MUSCLE (Edgar, 2004) and was subsequently trimmed using Gblocks (Castresana, 2000). The resulting alignment was analyzed using ProtTest (Abascal et al., 2005) to find the most appropriate substitution model which in this case was the LG model. RAxML (Stamatakis, 2014) was used to generate a maximum-likelihood phylogeny for this protein family and the resulting tree was viewed using Dendroscope (Huson et al., 2007) to identify the location of YedY sequences of interest. All known YedY sequences mapped to one of two clades which we named the "housekeeping" and "HGT" clades. All taxa from Figure 12, the species mnemonic for each proteome, and the locus tags of the associated genes are listed in Table 4. The taxonomic affiliation of each sequence used to generate the distribution data in Figure 12 are also included in Table 5. To generate the phylogeny shown in Figure 12, we concatenated the protein sequences from the adjacent *yedZ* gene to the HGT YedY sequences and a handful of housekeeping YedY genes to serve as the outgroup. The tree was then generated as described above except using the WAG+F model and a specific outgroup in RAxML. We performed ten independent inferences with RAxML, and selected the most likely tree. 100 bootstrap replicates were also performed and node support values were drawn onto the most likely tree.

References

Abascal, F., Zardoya, R., and Posada, D. (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics (Oxford, England)* **21**: 2104–2105.

Alvarez-Martinez, C.E., Baldini, R.L., and Gomes, S.L. (2006) A caulobacter crescentus extracytoplasmic function sigma factor mediating the response to oxidative stress in stationary phase. *J Bacteriol* **188**: 1835–1846.

Ames, G.F., Prody, C., and Kustu, S. (1984) Simple, rapid, and quantitative release of periplasmic proteins by chloroform. *J Bacteriol* **160**: 1181–1183.

Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.* **11**: R106.

Barloy-Hubler, F., Chéron, A., Hellégouarch, A., and Galibert, F. (2004) Smc01944, a secreted peroxidase induced by oxidative stresses in Sinorhizobium meliloti 1021. *Microbiology (Reading, England)* **150**: 657–664.

Bellaire, B.H., Elzer, P.H., Hagius, S., Walker, J., Baldwin, C.L., and Roop, R.M. (2003) Genetic organization and iron-responsive regulation of the Brucella abortus 2,3-dihydroxybenzoic acid biosynthesis operon, a cluster of genes required for wild-type virulence in pregnant cattle. *Infect. Immun.* **71**: 1794–1803.

Brokx, S.J., Rothery, R.A., Zhang, G., Ng, D.P., and Weiner, J.H. (2005) Characterization of an Escherichia coli sulfite oxidase homologue reveals the role of a conserved active site cysteine in assembly and function. *Biochemistry* **44**: 10339–10348.

Brzuszkiewicz, E., Brüggemann, H., Liesegang, H., Emmerth, M., Olschläger, T., Nagy, G., et al. (2006) How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic Escherichia coli strains. *Proceedings of the National Academy of Sciences* **103**: 12879–12884.

Carlson, H.K., Kuehl, J.V., Hazra, A.B., Mullan, M.R., Iavarone, A.T., Engelbrektson, A., Price, M.N., Deutschbauer, A.M., Arkin, A.P., and Coates, J.D. (2014) Direct inhibition of the respiratory sulfate-reduction pathway by (per)chlorate and nitrate. Manuscript in preparation.

Cartron, M.L., Maddocks, S., Gillingham, P., Craven, C.J., and Andrews, S.C. (2006) Feo- -transport of ferrous iron into bacteria. *Biometals* **19**: 143–157.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17**: 540–552.

Clark, I.C., Melnyk, R.A., Iavarone, A.T., Novichkov, P.S., and Coates, J.D. (2014) Chlorate reduction in Shewanella algae ACDC is a recently acquired metabolism characterized by gene loss, suboptimal regulation, and oxidative stress. *Mol Microbiol.* Manuscript in press.

Coates, J. and Achenbach, L. (2004) Microbial perchlorate reduction: rocket-fueled metabolism. *Nature Reviews Microbiology* **2**: 569–580.

Crossman, L.C., Chaudhuri, R.R., Beatson, S.A., Wells, T.J., Desvaux, M., Cunningham, A.F., et al. (2010) A commensal gone bad: complete genome sequence of the

prototypical enterotoxigenic Escherichia coli strain H10407. *J Bacteriol* **192**: 5822–5831.

Dartigalongue, C., Missiakas, D., and Raina, S. (2001) Characterization of the Escherichia coli sigma E regulon. *The Journal of biological chemistry* **276**: 20866–20875.

Deutschbauer, A., Price, M.N., Wetmore, K.M., Shao, W., Baumohl, J.K., Xu, Z., et al. (2011) Evidence-Based Annotation of Gene Function in Shewanella oneidensis MR-1 Using Genome-Wide Fitness Profiling across 121 Conditions. *PLoS genetics* **7**: e1002385.

Dobrindt, U., Blum-Oehler, G., Nagy, G., Schneider, G., Johann, A., Gottschalk, G., and Hacker, J. (2002) Genetic structure and distribution of four pathogenicity islands (PAI I(536) to PAI IV(536)) of uropathogenic Escherichia coli strain 536. *Infect. Immun.* **70**: 6365–6372.

Drazic, A., Miura, H., Peschek, J., Le, Y., Bach, N.C., Kriehuber, T., and Winter, J. (2013) Methionine oxidation activates a transcription factor in response to oxidative stress. *Proceedings of the National Academy of Sciences of the United States of America*.

Dunwell, J.M., Purvis, A., and Khuri, S. (2004) Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**: 7–17.

Eddy, S. (1998) Profile hidden Markov models. *Bioinformatics (Oxford, England)* **14**: 755–763.

Edgar, R. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* **5**: 113.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences* **95**: 14863–14868.

Ezraty, B., Aussel, L., and Barras, F. (2005) Methionine sulfoxide reductases in prokaryotes. *Biochim Biophys Acta* **1703**: 221–229.

Finotello, F., Lavezzo, E., Bianco, L., Barzon, L., Mazzon, P., Fontana, P., et al. (2014) Reducing bias in RNA sequencing data: a novel approach to compute counts. *BMC bioinformatics* **15 Suppl 1**: S7.

Gebendorfer, K.M., Drazic, A., Le, Y., Gundlach, J., Bepperling, A., Kastenmüller, A., et al. (2012) Identification of a hypochlorite-specific transcription factor from Escherichia coli. *The Journal of biological chemistry* **287**: 6892–6903.

Gray, M.J., Wholey, W.-Y., and Jakob, U. (2013) Bacterial Responses to Reactive Chlorine Species. *Annual Review of Microbiology*.

Gray, M.J., Wholey, W.-Y., Parker, B.W., Kim, M., and Jakob, U. (2013) NemR is a Bleach-Sensing Transcription Factor. *The Journal of biological chemistry*.

Guyer, D.M., Kao, J.S., and Mobley, H.L. (1998) Genomic analysis of a pathogenicity island in uropathogenic Escherichia coli CFT073: distribution of homologous sequences among isolates from patients with pyelonephritis, cystitis, and Catheter-

associated bacteriuria and from fecal samples. *Infect. Immun.* **66**: 4411–4417.
Ha, E.-M., Oh, C.-T., Bae, Y.S., and Lee, W.-J. (2005) A direct role for dual oxidase in Drosophila gut immunity. *Science (New York, N.Y.)* **310**: 847–850.
Havelius, K.G.V., Reschke, S., Horn, S., Döring, A., Niks, D., Hille, R., et al. (2011) Structure of the molybdenum site in YedY, a sulfite oxidase homologue from Escherichia coli. *Inorg Chem* **50**: 741–748.
Hawkins, C.L., Pattison, D.I., and Davies, M.J. (2003) Hypochlorite-induced oxidation of amino acids, peptides and proteins. *Amino Acids* **25**: 259–274.
Hofbauer, S., Gruber, C., Pirker, K.F., Sündermann, A., Schaffner, I., Jakopitsch, C., et al. (2014) Transiently produced hypochlorite is responsible for the irreversible inhibition of chlorite dismutase. *Biochemistry*.
Hull, R.A., Rudy, D.C., Donovan, W.H., Wieser, I.E., Stewart, C., and Darouiche, R.O. (1999) Virulence properties of Escherichia coli 83972, a prototype strain associated with asymptomatic bacteriuria. *Infect. Immun.* **67**: 429–432.
Huson, D., Richter, D., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. (2007) Dendroscope: An interactive viewer for large phylogenetic trees. *BMC bioinformatics* **8**: 460.
Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., et al. (2002) Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110. *DNA Res.* **9**: 189–197.
Kengen, S.W.S., Rikken, G.B.G., Hagen, W.R.W., van Ginkel, C.G.C., and Stams, A.J.A. (1999) Purification and characterization of (per)chlorate reductase from the chlorate-respiring strain GR-1. *J Bacteriol* **181**: 6706–6711.
Kohler, C., Lourenço, R.F., Avelar, G.M., and Gomes, S.L. (2012) Extracytoplasmic function (ECF) sigma factor sigmaF is involved in Caulobacter crescentus response to heavy metal stress. *BMC Microbiol* **12**: 210.
Kwolek-Mirek, M., Bartosz, G., and Spickett, C.M. (2011) Sensitivity of antioxidant-deficient yeast to hypochlorite and chlorite. *Yeast* **28**: 595–609.
Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**: 357–359.
Le, D.T., Liang, X., Fomenko, D.E., Raza, A.S., Chong, C.-K., Carlson, B.A., et al. (2008) Analysis of methionine/selenomethionine oxidation and methionine sulfoxide reductase function using methionine-rich proteins and antibodies against their oxidized forms. *Biochemistry* **47**: 6685–6694.
Lee, A.Q., Streit, B.R., Zdilla, M.J., Abu-Omar, M.M., and Dubois, J.L. (2008) Mechanism of and exquisite selectivity for O-O bond formation by the heme-dependent chlorite dismutase. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 15654–15659.
Levine, R.L., Mosoni, L., Berlett, B.S., and Stadtman, E.R. (1996) Methionine residues as

endogenous antioxidants in proteins. *Proceedings of the National Academy of Sciences* **93**: 15036–15040.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**: 2078–2079.

Li, W. and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)* **22**: 1658–1659.

Lima, T., Auchincloss, A., Coudert, E., Keller, G., Michoud, K., Rivoire, C., et al. (2009) HAMAP: a database of completely sequenced microbial proteome sets and manually curated microbial protein families in UniProtKB/Swiss-Prot. *Nucleic acids research* **37**: D471–8.

Lloyd, A.L., Rasko, D.A., and Mobley, H.L.T. (2007) Defining genomic islands and uropathogen-specific genes in uropathogenic Escherichia coli. *J Bacteriol* **189**: 3532–3546.

Loschi, L., Brokx, S.J., Hills, T.L., Zhang, G., Bertero, M.G., Lovering, A.L., et al. (2004) Structural and biochemical identification of a novel bacterial oxidoreductase. *The Journal of biological chemistry* **279**: 50391–50400.

López-Martínez, G., Rodríguez-Porrata, B., Margalef-Català, M., and Cordero-Otero, R. (2012) The STF2p hydrophilin from Saccharomyces cerevisiae is required for dehydration stress tolerance. *PLoS One* **7**: e33324.

Masloboeva, N., Reutimann, L., Stiefel, P., Follador, R., Leimer, N., Hennecke, H., et al. (2012) Reactive Oxygen Species-Inducible ECF σ Factors of Bradyrhizobium japonicum. *PLoS One* **7**: e43421.

Melnyk, R.A., Clark, I.C., Liao, A., and Coates, J.D. (2013) Transposon and deletion mutagenesis of genes involved in perchlorate reduction in Azospira suillum PS. *mBio* **5**: e00769–13.

Melnyk, R.A., Engelbrektson, A., Clark, I.C., Carlson, H.K., Byrne-Bailey, K., and Coates, J.D. (2011) Identification of a perchlorate reduction genomic island with novel regulatory and metabolic genes. *Appl Environ Microbiol* **77**: 7401–7404.

Oh, J., Fung, E., Price, M.N., Dehal, P.S., Davis, R.W., Giaever, G., et al. (2010) A universal TagModule collection for parallel genetic analysis of microorganisms. *Nucleic acids research* **38**: e146.

Oshlack, A. and Wakefield, M.J. (2009) Transcript length bias in RNA-seq data confounds systems biology. *Biol. Direct* **4**: 14.

Pan, B., Abel, J., Ricci, M.S., Brems, D.N., Wang, D.I.C., and Trout, B.L. (2006) Comparative oxidation studies of methionine residues reflect a structural effect on chemical kinetics in rhG-CSF. *Biochemistry* **45**: 15430–15443.

Pattison, D.I. and Davies, M.J. (2001) Absolute rate constants for the reaction of

hypochlorous acid with protein side chains and peptide bonds. *Chem. Res. Toxicol.* **14**: 1453–1464.
Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* **29**: e45.
Risso, D., Schwartz, K., Sherlock, G., and Dudoit, S. (2011) GC-content normalization for RNA-Seq data. *BMC bioinformatics* **12**: 480.
Roberts, A. and Pachter, L. (2013) Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat. Methods* **10**: 71–73.
Rodríguez-Porrata, B., Carmona-Gutierrez, D., Reisenbichler, A., Bauer, M., Lopez, G., Escoté, X., et al. (2012) Sip18 hydrophilin prevents yeast cell death during desiccation stress. *J Appl Microbiol* **112**: 512–525.
Roop, R.M., Gaines, J.M., Anderson, E.S., Caswell, C.C., and Martin, D.W. (2009) Survival of the fittest: how Brucella strains adapt to their intracellular niche in the host. *Med. Microbiol. Immunol.* **198**: 221–238.
Rosen, H., Klebanoff, S.J., Wang, Y., Brot, N., Heinecke, J.W., and Fu, X. (2009) Methionine oxidation contributes to bacterial killing by the myeloperoxidase system of neutrophils. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 18686–18691.
Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003) TM4: a free, open-source system for microarray data management and analysis. *BioTechniques* **34**: 374–378.
Santos, R., Hérouart, D., Sigaud, S., Touati, D., and Puppo, A. (2001) Oxidative burst in alfalfa-Sinorhizobium meliloti symbiotic interaction. *Mol. Plant Microbe Interact.* **14**: 86–89.
Schneider, G., Dobrindt, U., Middendorf, B., Hochhut, B., Szijártó, V., Emody, L., and Hacker, J. (2011) Mobilisation and remobilisation of a large archetypal pathogenicity island of uropathogenic Escherichia coli in vitro support the role of conjugation for horizontal transfer of genomic islands. *BMC Microbiol* **11**: 210.
Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* **30**: 1312–1313.
Staroń, A., Sofia, H.J., Dietrich, S., Ulrich, L.E., Liesegang, H., and Mascher, T. (2009) The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol Microbiol* **74**: 557–581.
Tarrago, L., Kaya, A., Weerapana, E., Marino, S.M., and Gladyshev, V.N. (2012) Methionine sulfoxide reductases preferentially reduce unfolded oxidized proteins and protect cells from oxidative protein unfolding. *The Journal of biological chemistry* **287**: 24448–24459.
van Ginkel, C.G., Rikken, G.B., Kroon, A.G.M., and Kengen, S.W.M. (1996) Purification

and characterization of chlorite dismutase: a novel oxygen-generating enzyme. *Arch Microbiol* **166**: 321–326.

Wang, S., Deng, K., Zaremba, S., Deng, X., Lin, C., Wang, Q., et al. (2009) Transcriptomic response of Escherichia coli O157:H7 to oxidative stress. *Appl Environ Microbiol* **75**: 6110–6123.

Wattam, A.R., Foster, J.T., Mane, S.P., Beckstrom-Sternberg, S.M., Beckstrom-Sternberg, J.M., Dickerman, A.W., et al. (2013) Comparative phylogenomics and evolution of the brucellae: A path to virulence. *J Bacteriol*.

Workun, G.J., Moquin, K., Rothery, R.A., and Weiner, J.H. (2008) Evolutionary persistence of the molybdopyranopterin-containing sulfite oxidase protein fold. *Microbiol Mol Biol Rev* **72**: 228–48.

Tables and Figures

Figure 1. *Chlorite stress phenotypes of Δ*sigF *and Δ*nrsF. (**A**) Impact of chlorite treatment during log phase growth on ALP media. (**B**) 160 µM chlorite is added immediately after inoculation during lag phase.

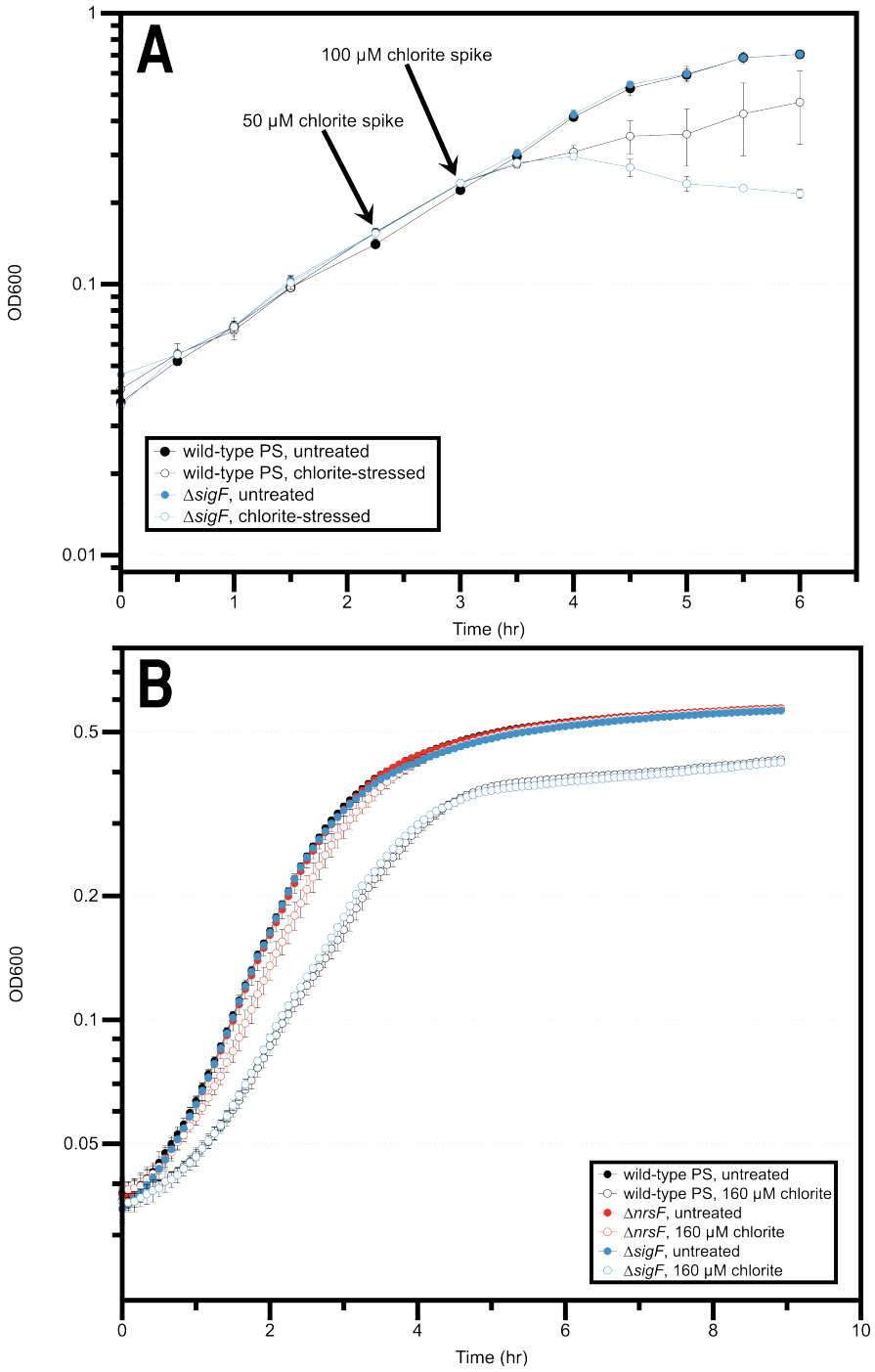


Figure 2. *Anaerobic growth of Δ*sigF *and Δ*nrsF *on chlorate.* The two mutant strains, as well as wild-type PS, were grown on minimal media containing 30 mM lactate and two different concentrations of chlorate.

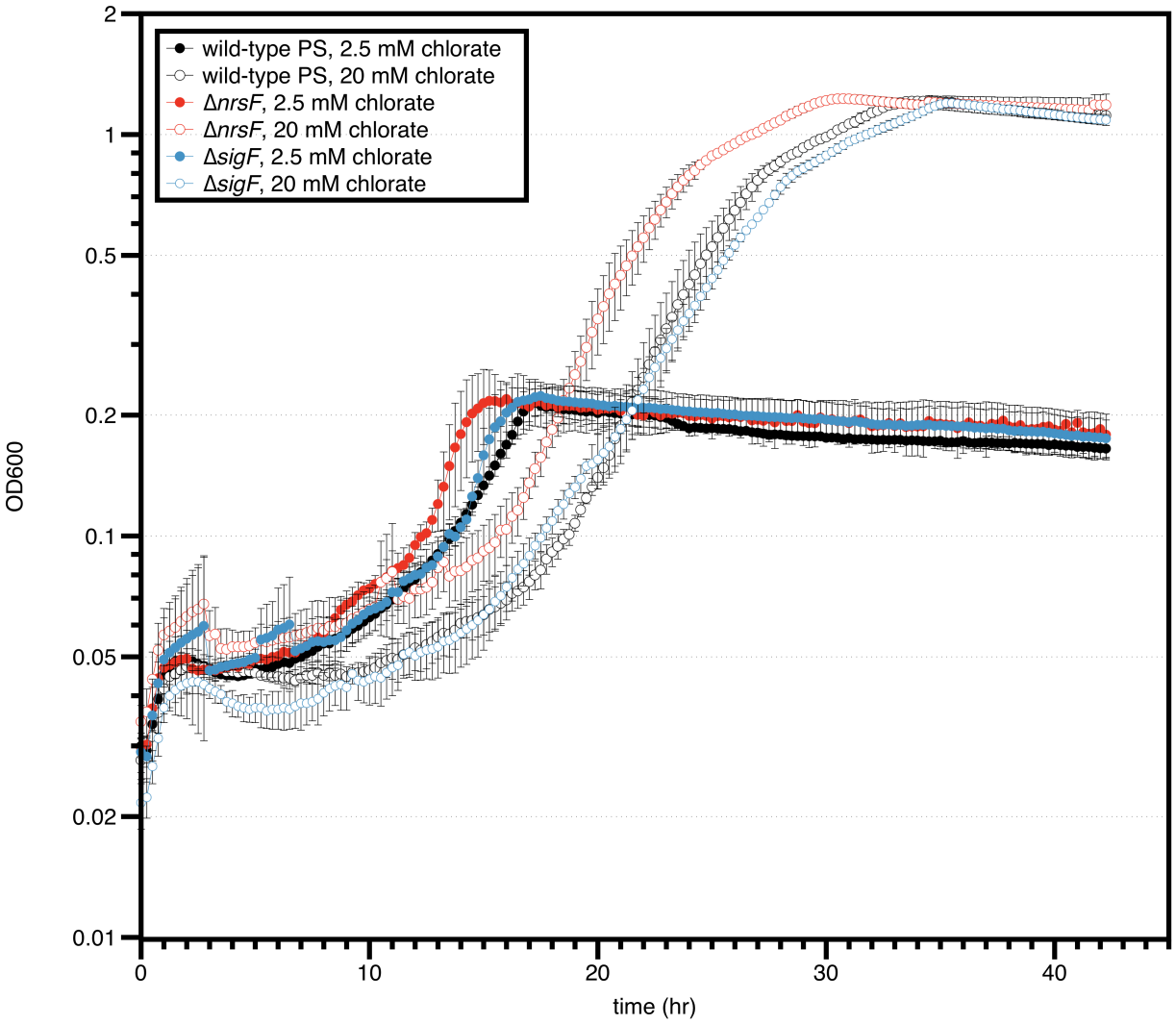


Figure 3. *Diagram of the SigF regulon, coverage, and promoter structure.* Average per-base coverage (left y-axis) across all RNA-seq replicates in the region of the *sigFnrsF* and *yedY1Z1mrpX* operons. The inset shows the identified transcriptional start sites and the SigF promoter found upstream of the start sites. The nucleotides also conserved in *Caulobacter crescentus* and *Bradyrhizobium japonicum* are indicated in bold. The purple dots show the GC-content of the chromosome over a sliding 100-bp window and are intended to illustrate that the *mrpX* gene has a biased nucleotide composition even compared to the rest of the PRI.

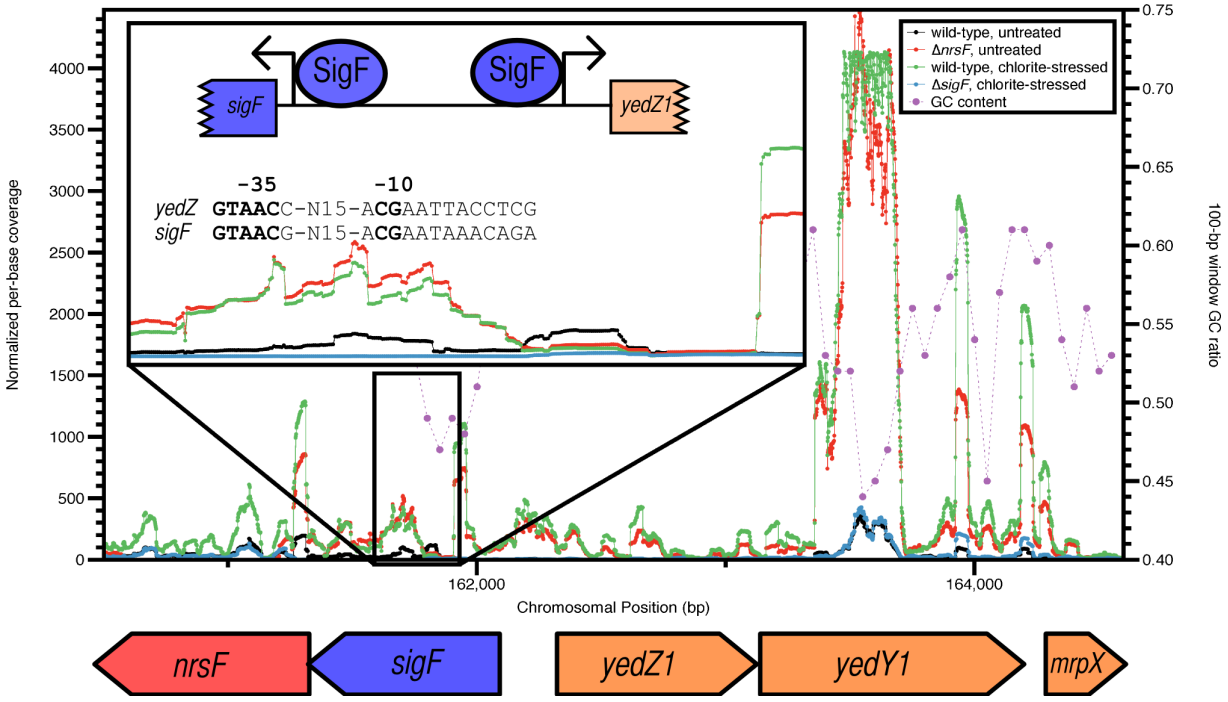
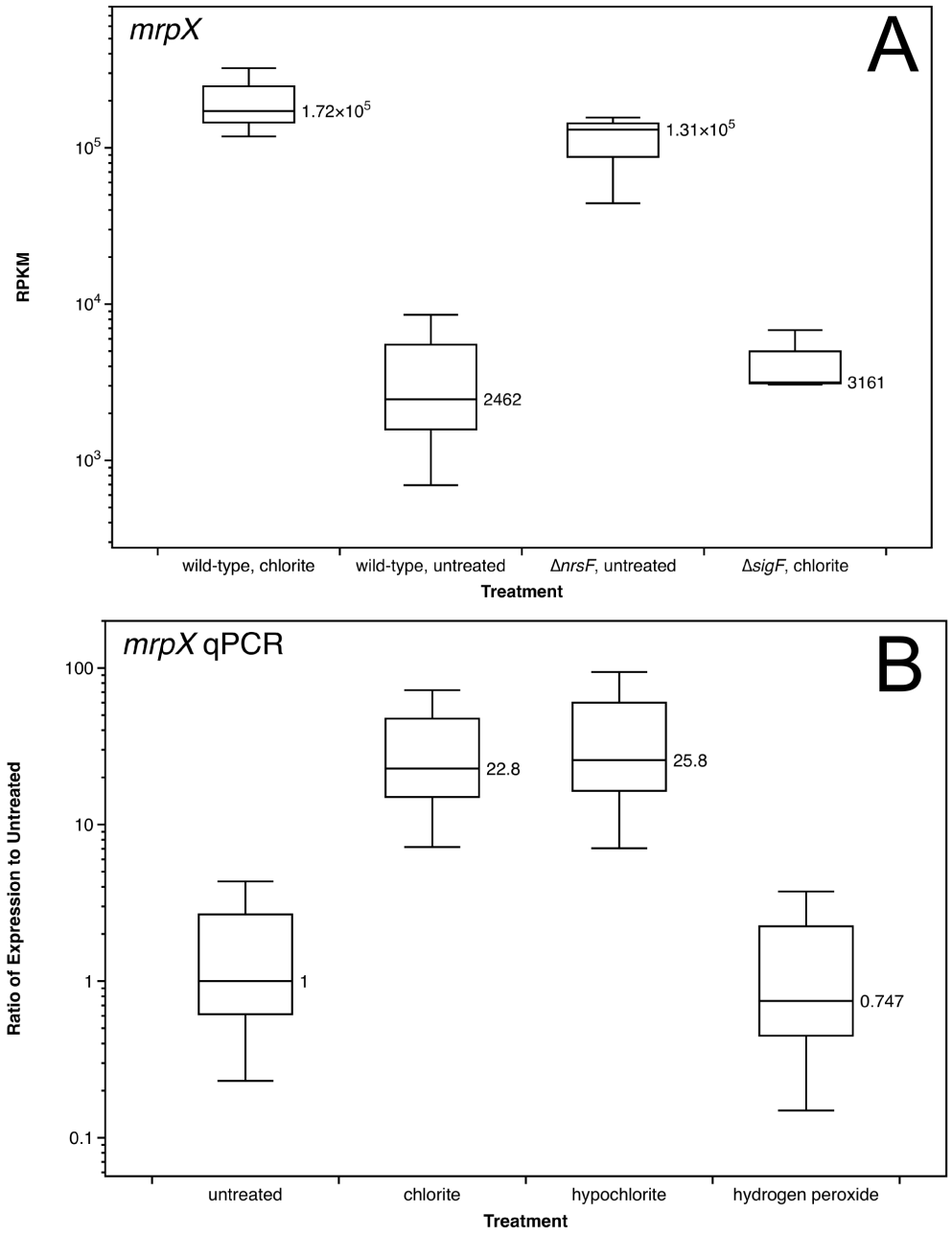
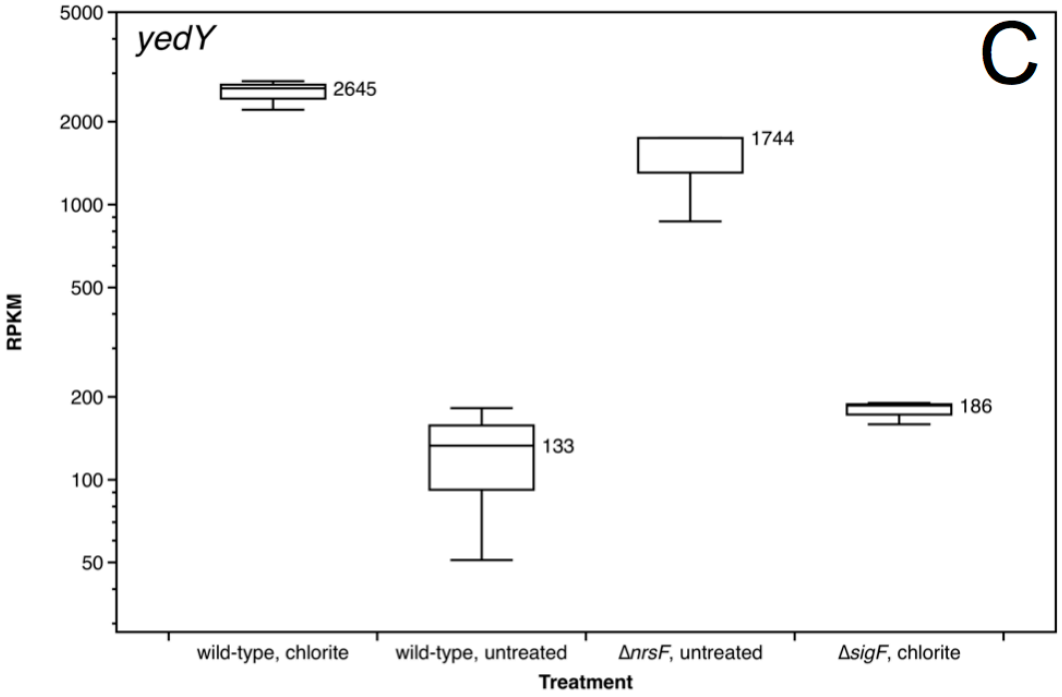


Figure 4. *Expression of genes in the SigF regulon.* Boxplots indicating the level of expression of *mrpX* (**A**, RNA-seq and **B**, qPCR), *yedY1* (**C**, RNA-seq), and *yedZ1* (**D**, RNA-seq). For the RNA-seq plots, the values are expressed in reads per kilobase per million mapped reads (RPKM), which is a value normalized by gene length and depth of coverage. The qPCR values are expressed as relative abundance to *mrpX* expression in the untreated control and the whiskers indicate standard error across three biological replicates.




yedY
C
5000
2000
1000
500
200
100
50
RPKM
2645
1744
133
186
wild-type, chlorite
wild-type, untreated
ΔnrsF, untreated
ΔsigF, chlorite
Treatment

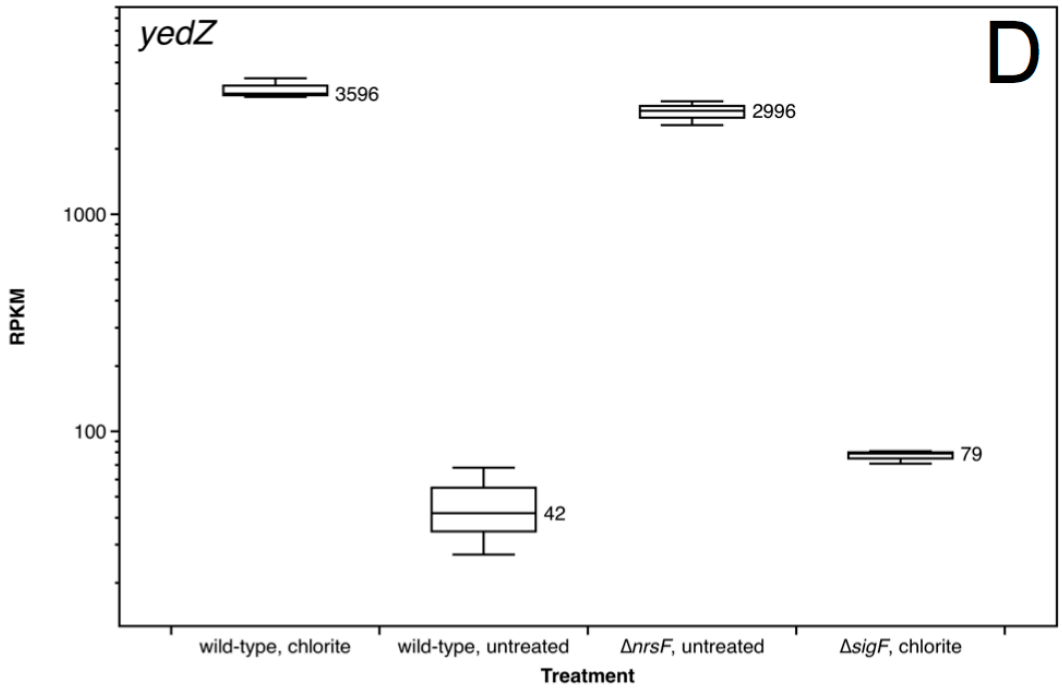

yedZ
D
1000
100
RPKM
3596
2996
42
79
wild-type, chlorite
wild-type, untreated
ΔnrsF, untreated
ΔsigF, chlorite
Treatment

Figure 5. *Expression of pcrA and cld.* Boxplots of *pcrA* (A) and *cld* (B) expression. Values are expressed in reads per kilobase per million mapped reads (RPKM), which is a value normalized by gene length and depth of coverage.

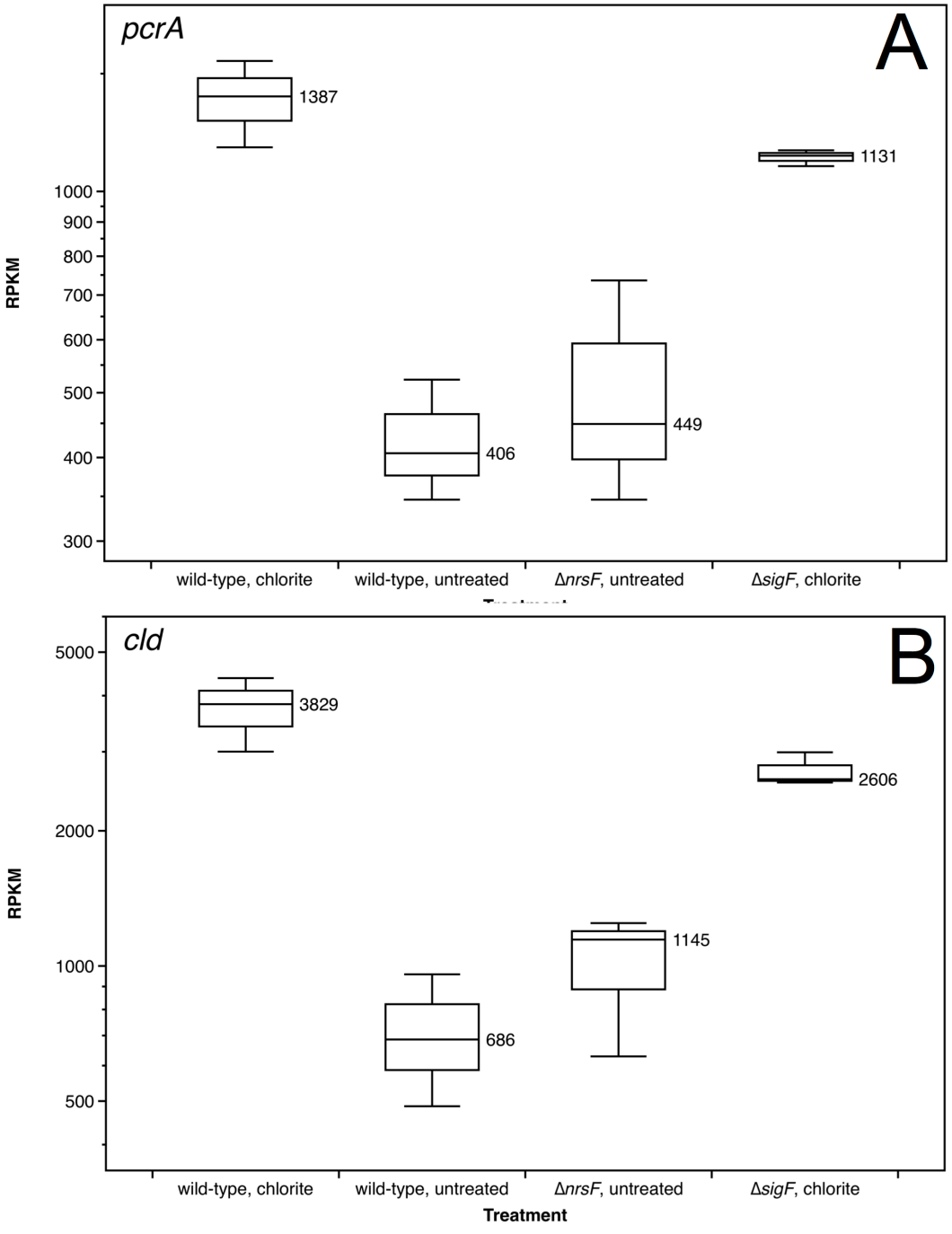


Figure 6. *Phenotypes of Δ*yedY1 *and* ΔmrpX *in the Δ*nrsF *background.* (**A**) Aerobic growth curve of wild-type, Δ*nrsF*, Δ*nrsF*Δ*yedY1*, and Δ*nrsF*Δ*mrpX* on ALP media either untreated or treated with 90 μM. (**B**) Anaerobic growth curve of the same four genotypes on minimal media with 30 mM lactate and either 2.5 or 20 mM.

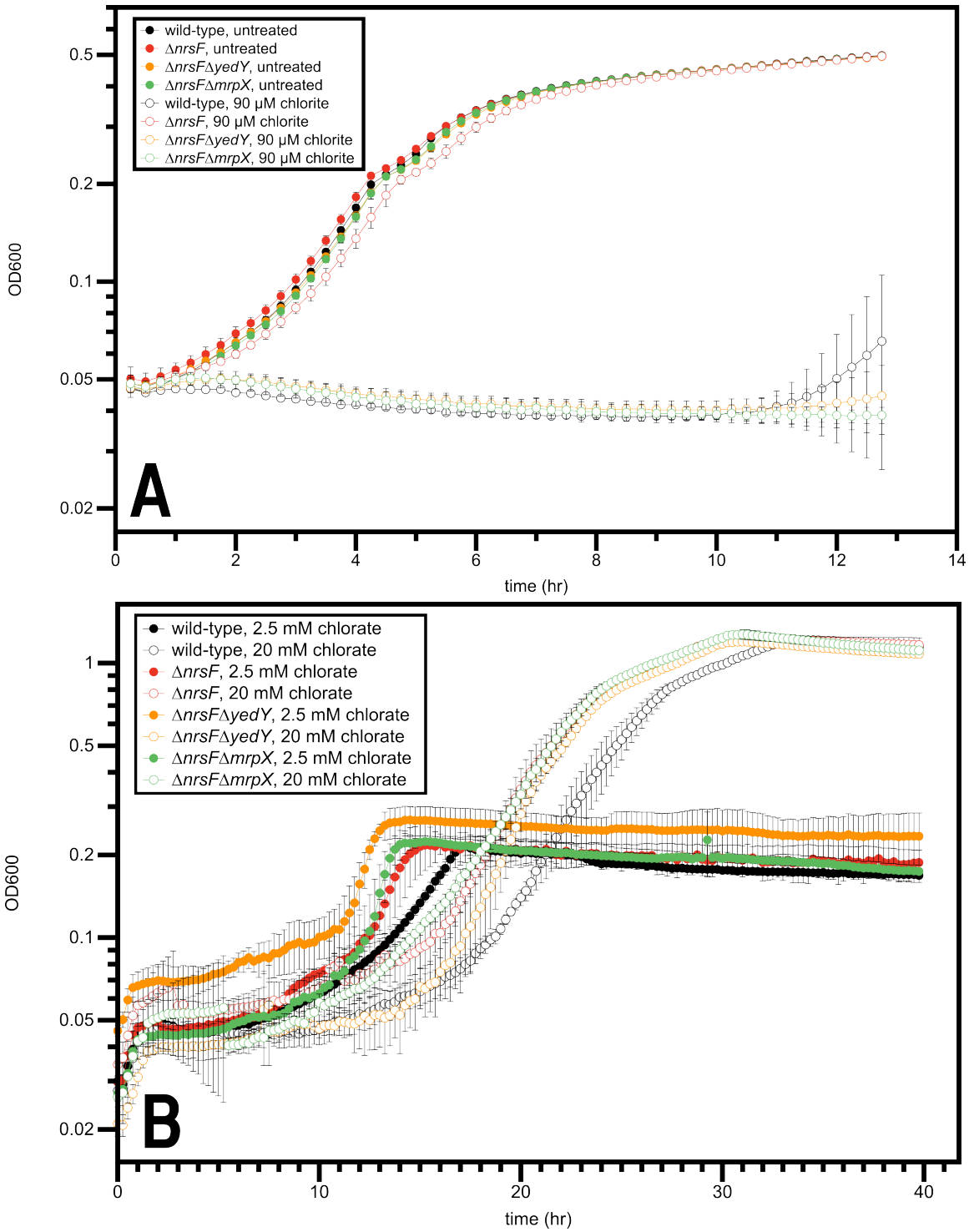


Figure 7. *BarSeq fitness data of relevant genes.* Heatmap generated using MeV to show the fitness values of (**A**) fitness of all 17 genes that make up the core of the PRI, (**B**) the cluster of genes that contain *yedY1Z1* and *yedY2Z2*, and (**C**) the cluster of genes with the strongest defects specific to the chlorite stress condition.

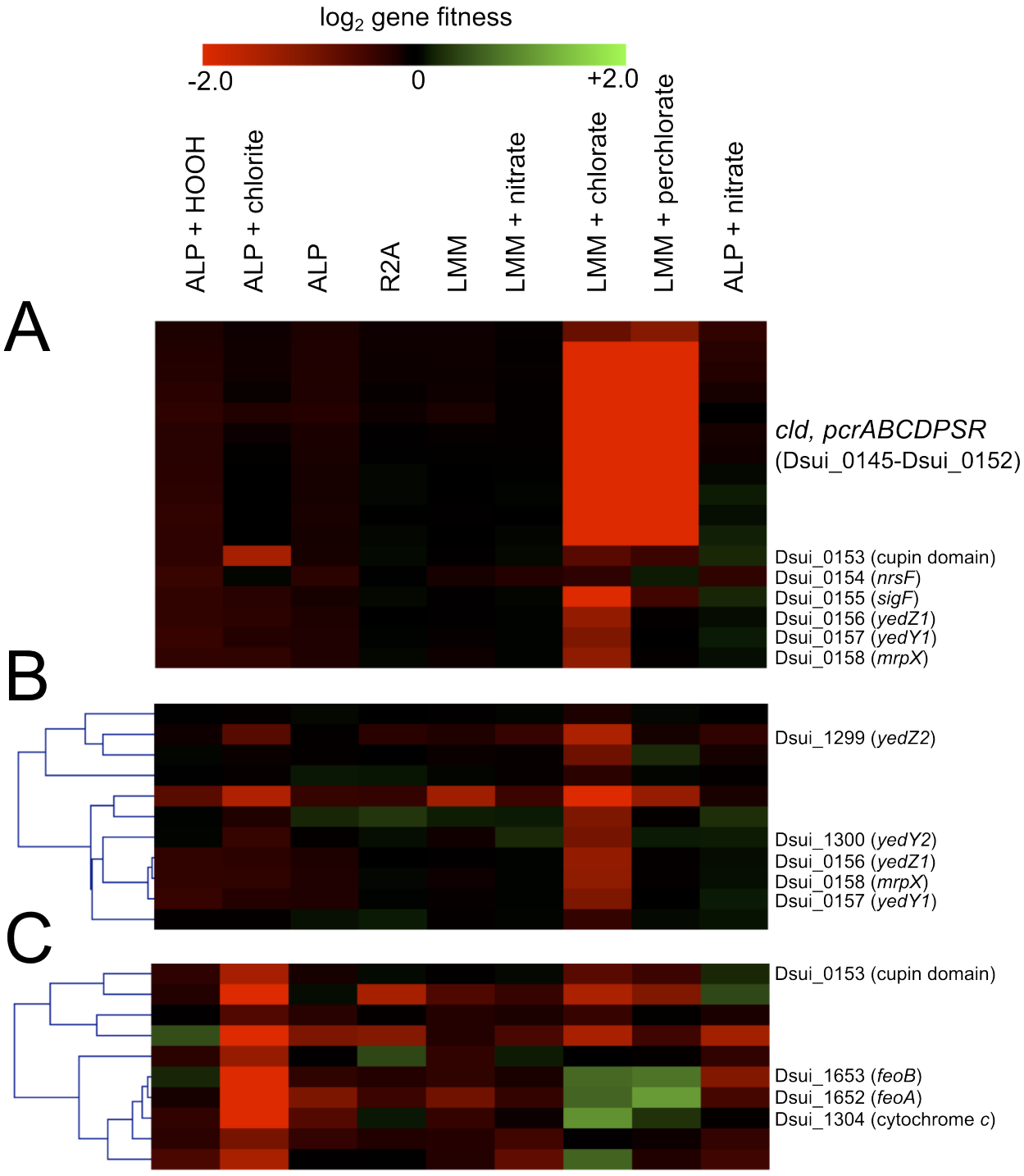


Figure 8. *Deletion of* yedY2 *reveals genetic redundancy.* Anaerobic growth curve showing the slight defect of the Δ*yedY1*Δ*yedY2* strain on 20 mM chlorate and severe defects of the Δ*sigF*Δ*yedY2* strain on both 2.5 mM and 20 mM chlorate. This experiment was carried out using minimal media containing 30 mM lactate with a wild-type control.

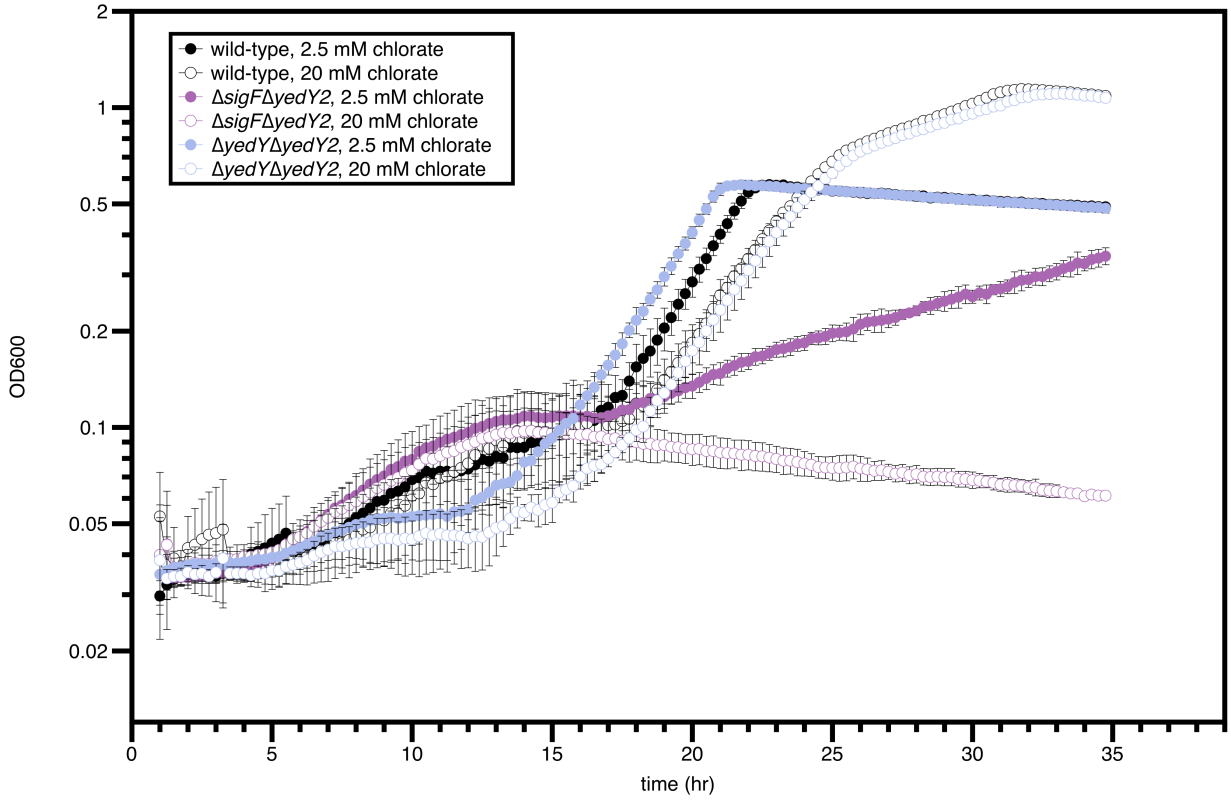


Figure 9. *Oxidation of MrpX detected via SDS-PAGE* (**A**) Total periplasmic protein and (**B**) western blot using an Anti-myc tag primary antibody and HRP-conjugated secondary. The samples run were from several genotypes: Δ*mrpX* (negative control), Δ*nrsF mrpX*::*myc* (positive control), and *mrpX*::*myc* (experiment).

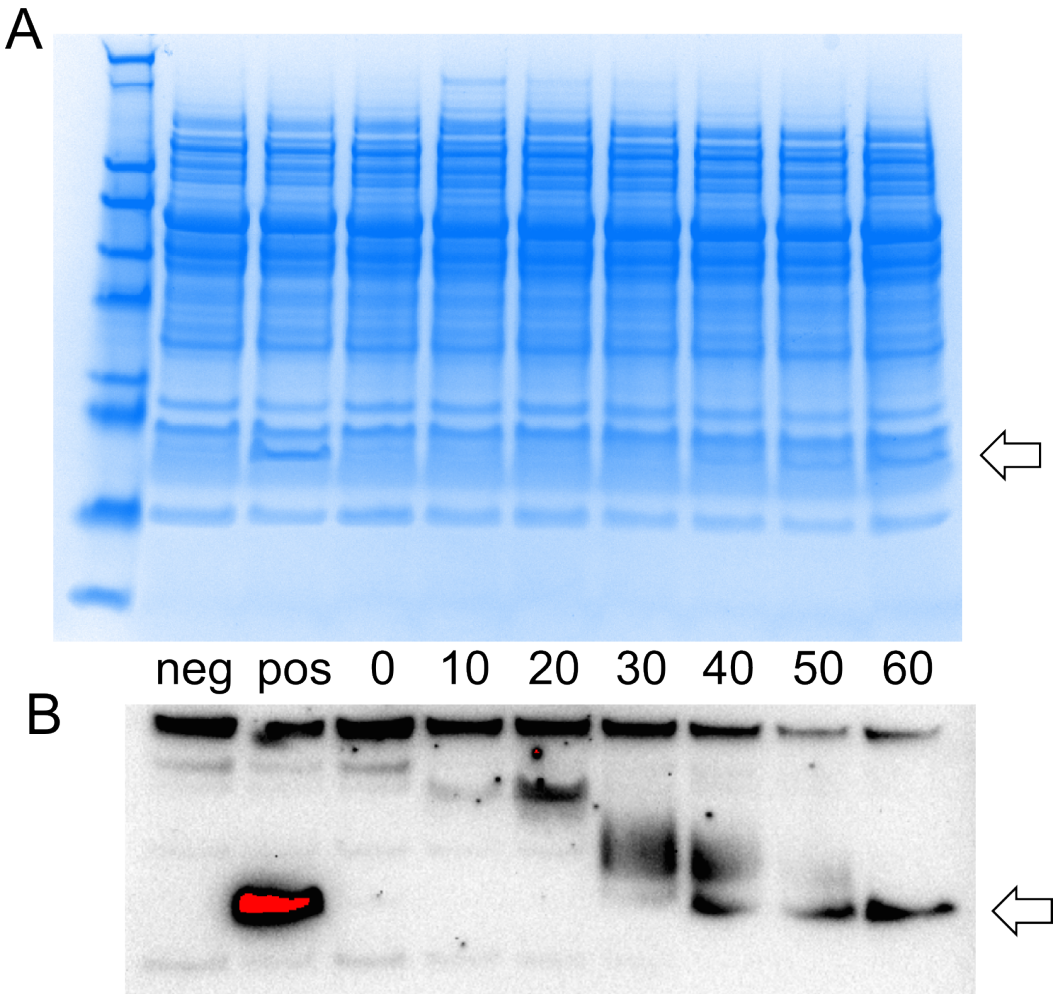


Figure 10. *Methionine oxidation of the global proteome.* Boxplot where the central line is the median of three biological replicates, while the whiskers represent the minimum and maximum. The y-axis measures the ratio of oxidized methionines detected to total methionine residues detected.

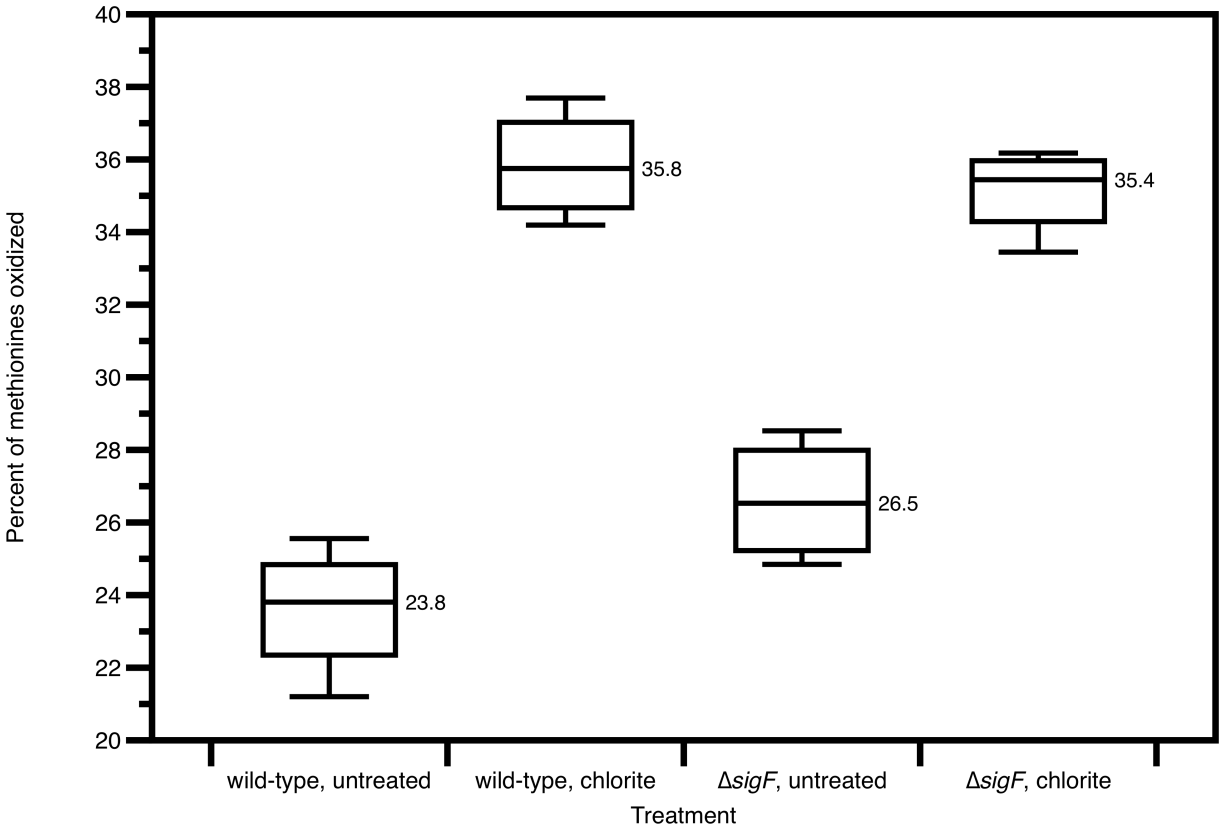


Figure 11. *Methionine shortens lag phase in a* yedY*-dependent manner.* Anaerobic growth curves showing addition of 200 μM methionine and an untreated control in minimal media containing 30 mM lactate for (**A**) wild-type with 2.5 mM perchlorate, (**B**) Δ*yedY*Δ*yedY2* with 2.5 mM perchlorate.

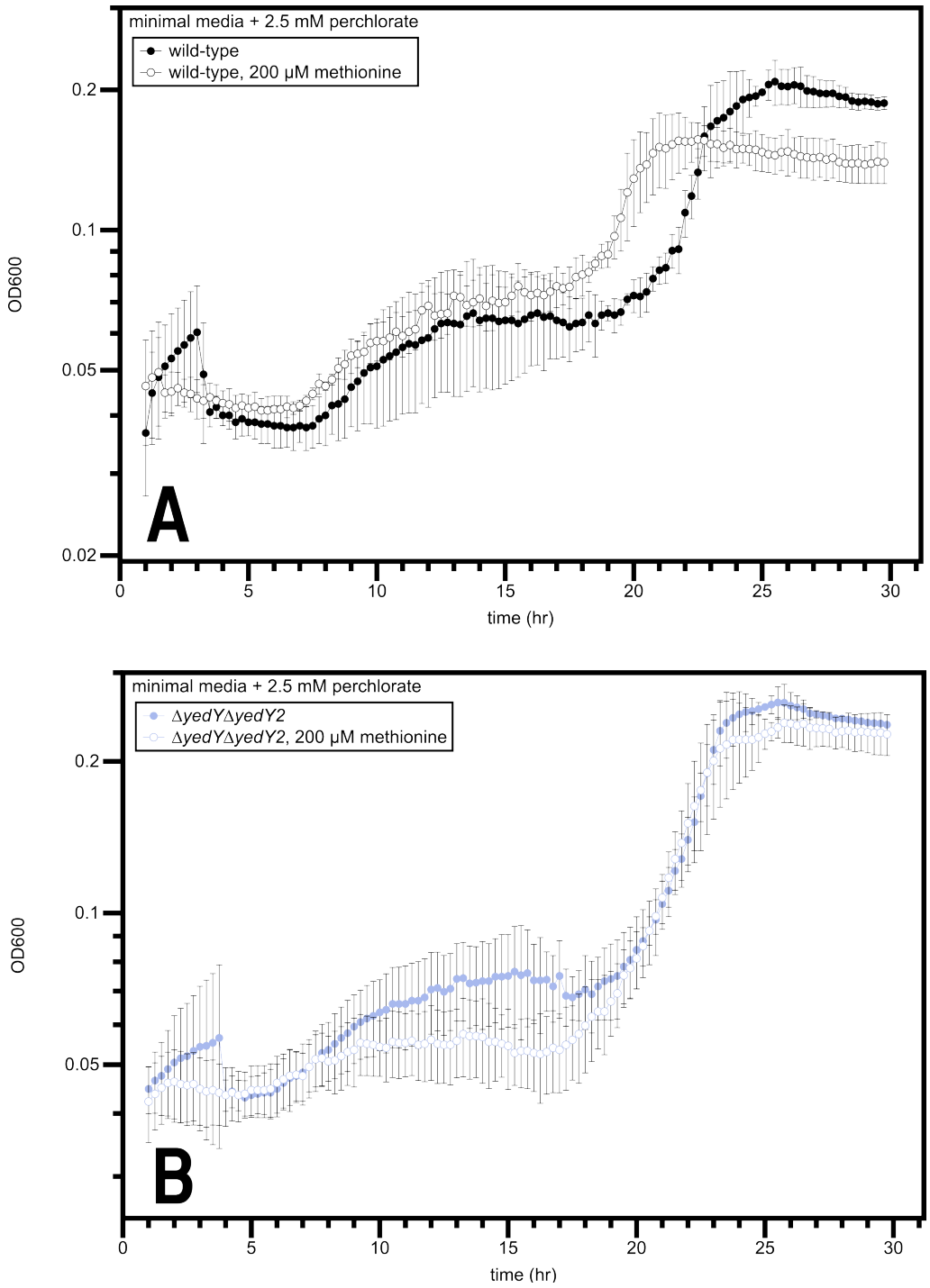


Figure 12. *Phylogeny of the 'HGT' clade of YedY*. Taxon labels show both the SwissProt ID and the species of origin. Colored star markers indicate the presence of conserved accessory genes nearby *yedY* in the genome. The gene diagrams at the left side of the tree show the striking difference in *yedYZ* synteny between the housekeeping and HGT clades. Bootstrap support for a given node is illustrated by black (> 90/100) or grey (> 70/100) circles.

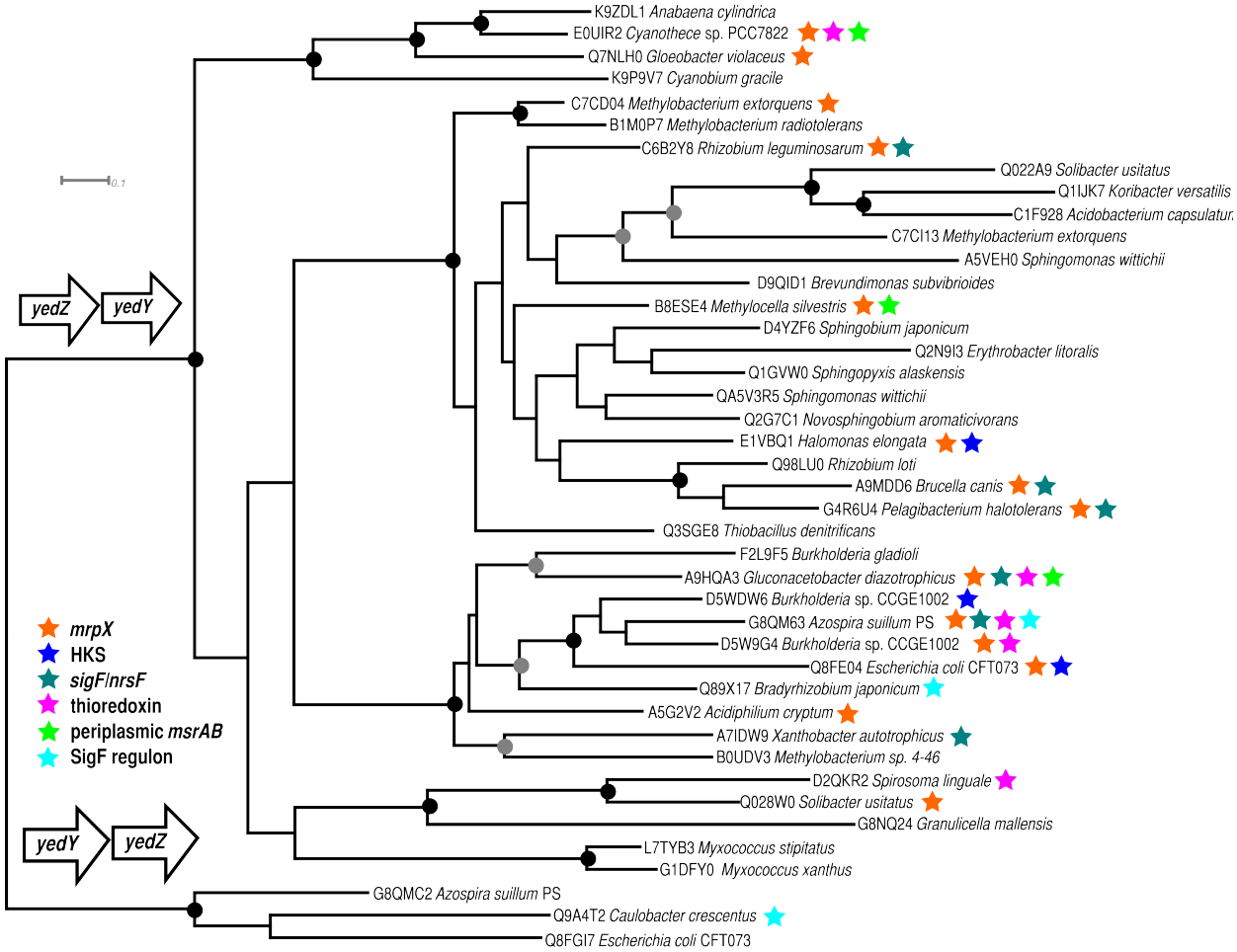


Figure 13. *Taxonomic distribution of the two clades of YedY.* Panel (**A**) details the distribution of YedY among the five major proteobacterial classes and panel (**B**) shows the distribution among several proteobacterial orders of interest.

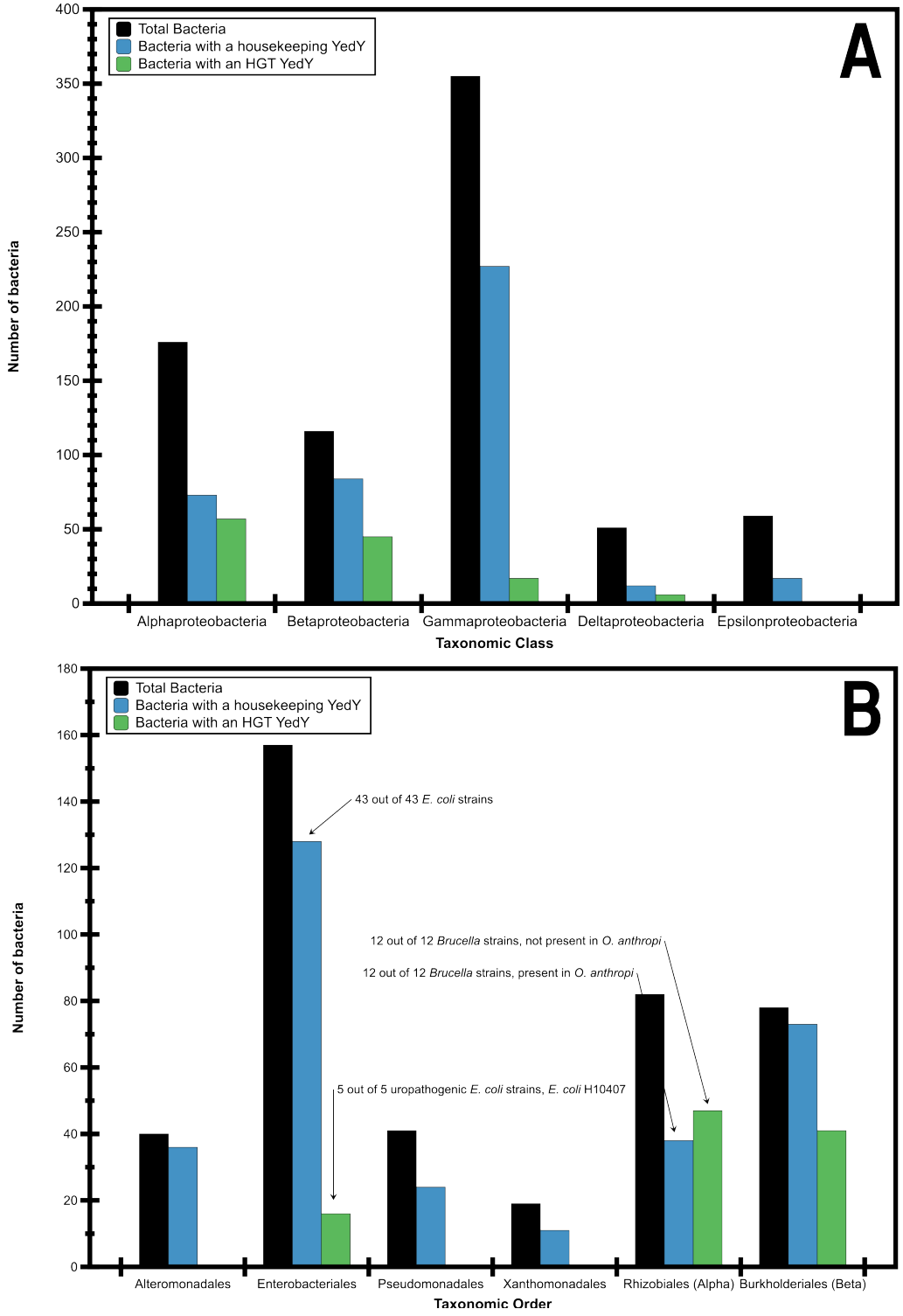


Figure 14. *Comparative genomic architecture of* yedYZ. This figure shows the genomic organization of genes of interest from several organisms, including perchlorate reducers and host-associated organisms.

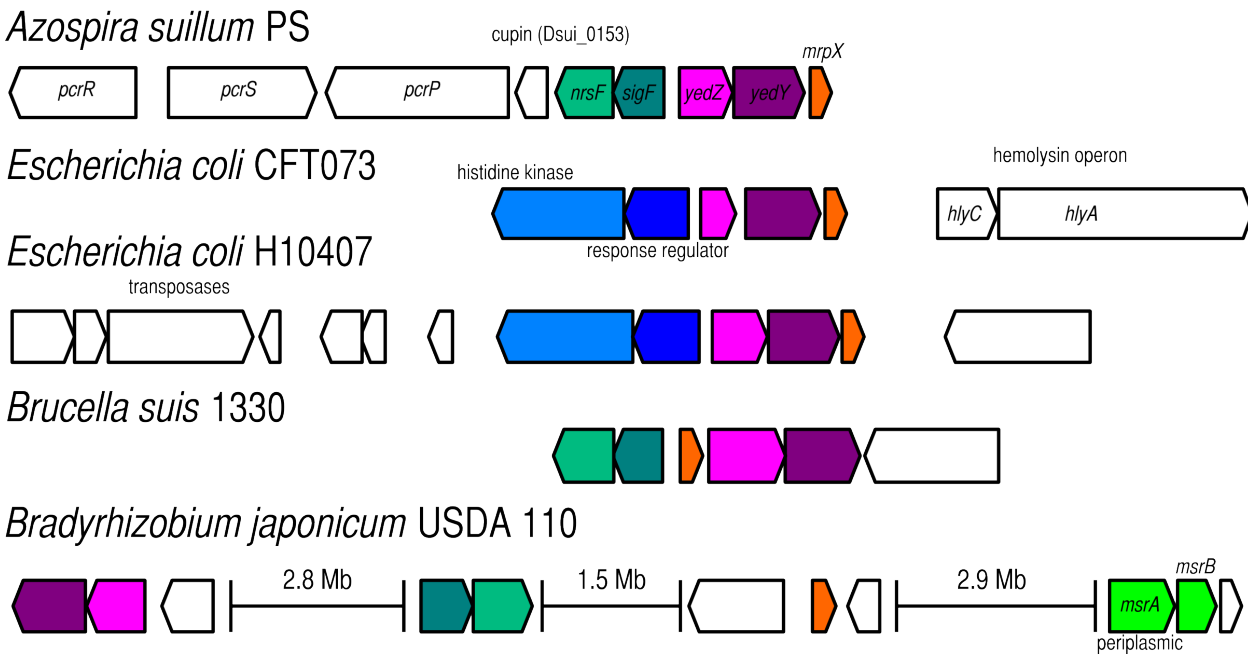


Table 1. *Strains used in this chapter.*

Strain Name	Genotype
RAM000	wild-type *Azospira suillum* PS
RAM026	ΔDsui_0154 (*nrsF*)
RAM027	ΔDsui_0157 (*yedY1*)
RAM028	ΔDsui_0155 (*sigF*)
RAM029	ΔDsui_0156 (*yedZ*)
RAM030	ΔDsui_0158 (*mrpX*)
RAM053	ΔDsui_0154 (*nrsF*) ΔDsui_0157 (*yedY1*)
RAM054	ΔDsui_0154 (*nrsF*) ΔDsui_0158 (*mrpX*)
RAM055	ΔDsui_0154 (*nrsF*) *mrpX*::*myc*
RAM056	*mrpX*::*myc*
RAM083	ΔDsui_1300 (*yedY2*)
RAM084	ΔDsui_0155 (*sigF*) ΔDsui_1300 (*yedY2*)
RAM085	ΔDsui_0157 (*yedY1*) ΔDsui_1300 (*yedY2*)

Table 2. *Plasmids used in this chapter.*

Plasmid	Derivative of	Marker(s)	Description
pRAM33	pNPTS138	Km, *sacB*	pNPTS138 with Dsui_0156 deletion insert
pRAM37	pNPTS138	Km, *sacB*	pNPTS138 with Dsui_0157 deletion insert
pRAM38	pNPTS138	Km, *sacB*	pNPTS138 with Dsui_0154 deletion insert
pRAM39	pNPTS138	Km, *sacB*	pNPTS138 with Dsui_0155 deletion insert
pRAM40	pNPTS138	Km, *sacB*	pNPTS138 with Dsui_0158 deletion insert
pRAM56	pCR BLUNT TOPO	Km	158comp a to d construct (2.1 kb)
pRAM59	pCR BLUNT TOPO	Km	158myc a to d construct in TOPO vector
pRAM60	pNPTS138	Km, *sacB*	pNPTS138 with Dsui_158 + myc tag
pRAM73	pNPTS138	Km, *sacB*	pNPTS138 with insert to delete Dsui_1300 (yedY2)

Table 3. *Primers used in this chapter.*

Primer Name	**Sequence**	**Description**
PS0158a	GNNACTAGTTGGCCATGGTCGCGTTGGTC	deletion of Dsui_0158 (*mrpX*)
PS0158d	GNNACTAGTAGGCTCGCTCCCTACTGGCA	deletion of Dsui_0158 (*mrpX*)
PS1300a	GNNGAATTCTTCACGAAGACGAGCTCCAG	deletion of Dsui_1300 (*yedY2*)
PS1300b	GNNGTCGACAATCTCGGATGGCAGCAGG	deletion of Dsui_1300 (*yedY2*)
PS1300c	GNNGTCGACCTCTACAGCGGCATGGACCT	deletion of Dsui_1300 (*yedY2*)
PS1300d	GNNGGATCCCTAGCGGTCAAGCCTTCCC	deletion of Dsui_1300 (*yedY2*)
PS0158+	CACCTCCTGCCTTATGGTCG	qPCR of *mrpX* (139 bp product)
PS0158-	GCCATCCTTGGCCATACTGT	qPCR of *mrpX* (139 bp product)
PSrpoB+	AGCAGGAAGTGTACATGGGC	qPCR of *rpoB* (Dsui_0366) 206 bp
PSrpoB-	TTCGAAGTCCAGCCAGGAAC	qPCR of *rpoB* (Dsui_0366) 206 bp

Table 4. *Locus tags of HGT YedY sequences.* These are the taxa from Figure 12.

SwissProt ID	Species Mnemonic	Locus Tag
Q8FE04	ECOL6	c3567
D5W9G4	BURSC	BC1002_1957
D5WDW6	BURSC	BC1002_4969
G8QM63	AZOSU	Dsui_0157
F2L9F5	BURGS	bgla_1g08620
A5VEH0	SPHWW	Swit_4346
Q028W0	SOLUE	Acid_1449
D2QKR2	SPILD	Slin_4244
G8NQ24	GRAMM	AciX8_4077
Q1DFY0	MYXXD	MXAN_0160
L7TYB3	MYXSD	MYSTI_00143
Q7NLH0	GLOVI	gll1153
E0UIR2	CYAP2	Cyan7822_1372
K9ZDL1	ANACC	Anacy_1808
K9P9V7	CYAGP	Cyagr_2655
Q1IJK7	KORVE	Acid345_3943
Q022A9	SOLUE	Acid_3217
C1F928	ACIC5	ACP_2091
C7CI13	METED	METDI4829
D9QID1	BRESC	Bresu_0119
Q2G7C1	NOVAD	Saro_1812
C6B2Y8	RHILS	Rleg_2568
Q2N9I3	ERYLH	ELI_07830
Q1GVW0	SPHAL	Sala_0491
A5V3R5	SPHWW	Swit_0563
Q98LU0	RHILO	mlr0879
A9MDD6	BRUC2	BCAN_B0020
G4R6U4	PELHB	KKY_2247
Q3SGE8	THIDA	Tbd_2349
D4YZF6	SPHJU	SJA_C1-09040
B8ESE4	METSB	Msil_0864
E1VBQ1	HALED	HELO_4324
B1M0P7	METRJ	Mrad2831_1008
C7CD04	METED	METDI4101
A7IDW9	XANP2	Xaut_0961
A5G2V2	ACICJ	Acry_2994
A9HQA3	GLUDA	GDI2760
B0UDV3	METS4	M446_4897
Q89X17	BRAJA	bll0505
G8QMC2	AZOSU	Dsui_1300
Q8FGI7	ECOL6	c2389
Q9A4T2	CAUCR	CC_2748

Table 5. *Abundance of* yedY *sequences in different taxonomic groups.*

Taxonomic Group	Housekeeping	HGT	Species
Caulobacterales (Alpha)	5	2	7
Hyphomonadales (Alpha)	0	0	3
Magnetococcales (Alpha)	1	0	1
Rhizobiales (Alpha)	38	47	82
Rhodobacterales (Alpha)	16	0	16
Rhodospirillales (Alpha)	11	3	12
Rickettsiales (Alpha)	0	0	43
SAR11 (Alpha)	1	0	2
SAR116 (Alpha)	1	0	1
Sphingomonadales (Alpha)	0	5	9
Burkholderiales (Beta)	73	41	78
Methylophilales (Beta)	4	0	5
Neisseriales (Beta)	3	1	20
Rhodocyclales (Beta)	4	2	6
Gallionellales (Beta)	0	0	2
Hydrogenophilales (Beta)	0	1	1
Nitrosomonadales (Beta)	0	0	4
Bdellovibrionales (Delta)	0	0	2
Desulfarculales (Delta)	0	1	1
Desulfobacterales (Delta)	0	0	8
Desulfovibrionales (Delta)	0	0	13
Desulfurellales (Delta)	0	0	1
Desulfuromonadales (Delta)	3	0	11
Myxococcales (Delta)	9	5	11
Syntrophobacterales (Delta)	0	0	4
Campylobacterales (Epsilon)	17	0	56
Nautiliales (Epsilon)	0	0	1
Nitratiruptor (Epsilon)	0	0	1
Sulfurovum (Epsilon)	0	0	1
Acidithiobacillales (Gamma)	0	0	3
Aeromonadales (Gamma)	4	0	5
Alteromonadales (Gamma)	36	0	40
Chromatiales (Gamma)	4	0	11
Enterobacteriales (Gamma)	128	16	157
Legionellales (Gamma)	0	0	11
Methylococcales (Gamma)	3	0	3
Oceanospirillales (Gamma)	4	1	9
Pasteurellales (Gamma)	9	0	22
Pseudomonadales (Gamma)	24	0	41
Thiotrichales (Gamma)	3	0	19

Vibrionales (Gamma)	1	0	15
Xanthomonadales (Gamma)	11	0	19
Actinobacteria	0	10	183
Acidobacteria	6	4	9
Bacteroidetes	1	1	79
Chloroflexi	7	0	16
Cyanobacteria	13	9	50
Deinococcus-Thermus	7	0	16
Planctomycetes	2	0	7
Gemmatimonadetes	1	0	1

Chapter 4:

The Perchlorate Reduction Genomic Island: Mechanisms and Pathways of Evolution by Horizontal Gene Transfer

Abstract

Dissimilatory perchlorate reduction is a two-step metabolic process in which the reduction of perchlorate to chloride is coupled to generation of a proton-motive force. The key genes responsible for this metabolism, *pcrA* and *cld*, are located on a genomic island in the model organism *Azospira suillum* PS called the perchlorate reduction genomic island (PRI). Here we report on 13 finished and draft genomes of perchlorate-reducing bacteria from four different classes of Phylum Proteobacteria (the Alpha-, Beta-, Gamma- and Epsilonproteobacteria). Among the different phylogenetic classes, the PRI varies considerably in genetic content as well as putative mechanism and location of integration. However, the PRIs of the densely sampled genera *Azospira* and *Magnetospirillum* have striking nucleotide identity despite divergent housekeeping genes, implying horizontal transfer and positive selection within narrow phylogenetic taxa. We also assess the phylogenetic origin of accessory genes in the various incarnations of the PRI, which can be traced to chromosomal homologs from phylogenetically similar organisms. These observations suggest a phylogenetic history where the PRI rarely is transferred at the class level but undergoes frequent and continuous recombination within narrow phylogenetic groups. This restricted HGT is seen directly by the independent integration of near-identical islands within a genus and indirectly due to the acquisition of lineage-specific accessory genes. This genomic reversibility of perchlorate reduction may present a unique equilibrium for a metabolism that confers a competitive advantage only in the presence of an electron acceptor which is sparsely distributed in nature.

Background

Perchlorate is a sparsely distributed anion that is toxic to humans, but serves as a valuable electron acceptor for several lineages of bacteria. These bacteria have been isolated from many different taxonomic groups, although they seem to be restricted to the Proteobacteria. The metabolism depends on the reduction of perchlorate to chlorite in the bacterial periplasm by perchlorate reductase, and the dismutation of chlorite to chloride and molecular oxygen. Molecular oxygen is then used as an electron acceptor, generally through the use of a high-affinity cytochrome *cbb*$_3$ oxidase; the microaerobic portion of this respiration means that dissimilatory perchlorate-reducing bacteria (PRB) uniquely straddle the realms of both aerobic and anaerobic respirations. Genetic studies in the model PRB *Azospira suillum* PS have revealed that many of the genes specifically essential for this metabolism are located on a genomic island called the **p**erchlorate **r**eduction genomic **i**sland (PRI) (Melnyk et al., 2013), which was previously identified in a comparative genomic analysis of four PRB genomes (Melnyk et al., 2011).

This earlier computational analysis of PRB made use of only four isolates: three of which were from the Betaproteobacteria. Since then, there has been isolation of bacteria from other phylogenetic groups, not to mention dozens of unsequenced isolates. In this chapter, we report on our analysis of 13 PRB genomes that our lab has sequenced, both draft and complete, from four taxonomic classes: Alpha-, Beta-, Gamma- and Epsilonproteobacteria. By looking at the genomic architecture of the PRI and the history of individual genes within the island, we can piece together a probable hypothesis for how this metabolism came to exist, and how it continues to evolve today. What we find is a metabolism that reflects a history of horizontal gene transfer, both ancient and recent. Our model provides a framework for understanding this metabolism from an evolutionary and ecological basis, as our field moves increasingly towards attempting to understand PRB in the natural environment.

Results

Perchlorate/chlorate reducers are diverse members of the Proteobacteria but are restricted to certain subclades

To assess the phylogenetic relationships between PRB and other bacteria, we adopted a multi-locus sequence analysis (MLSA) method that relies upon an in-house pipeline for identifying as many orthologs as possible within a given group of bacteria. While MLSA methods do not fully recapitulate the reticulate and piecewise nature of bacterial genome evolution and speciation (Retchless and Lawrence, 2010), they are generally satisfactory for resolving a consensus pattern of species divergence (Hanage

et al., 2005), particularly when many "taxonomically complete" loci are used (Bonaventura et al., 2010).

Our ortholog detection pathway was based on a database of pairwise InParanoid analyses of a large group of bacteria (Remm et al., 2001). This computationally intensive step was then followed by rapid QuickParanoid clustering of orthologous sequences (http://pl.postech.ac.kr/QuickParanoid/). Using a very rapid clustering step enabled us to test the effects of taxon deletion on ortholog clustering, as well as rapidly generate clusters for varying taxonomic levels. To this end, we generated protein matrices for proteomes from all Proteobacteria (61 taxa, 92 genes; Figure 1), the Alphaproteobacteria (97 taxa, 80 genes; Figure 2), Betaproteobacteria (87 taxa, 147 genes; Figure 3), and Rhodocyclaceae (19 taxa, 229 genes; Figure 4) and constructed maximum likelihood trees based on these matrices (Figures 1-4).

While PRB are scatterered throughout the Proteobacteria (Figure 1), their presence within classes seems to be restricted. In the Alphaproteobacteria, PRB can only be found in the family Rhodospirillaceae (Figure 2), while in the Betaproteobacteria, PRB are only present in one subgroup within the Rhodocyclaceae (Figures 3, 4). This has been observed since the advent of molecular phylogeny in the context of PRB (Coates and Achenbach, 2004), but these phylogenetic inferences make use of the entire genome, rather than the 16S genes of PRB isolates (Coates et al., 1999; Thrash et al., 2010). PRB isolates from the Epsilonproteobacteria and Gammaproteobacteria have only recently been characterized and sequenced (Carlström et al., 2013), thus it is impossible to make sense of their distribution within their respective taxonomic classes until more strains have been isolated.

PRIs have a diverse, yet functionally similar, 'core' set of genes

The original identification of the PRI consisted of 4 separate island 'cores', which were defined by beginning at the *pcrA* and *cld* genes and using pairwise BLAST analysis to detect genes shared by more than one PRB (Melnyk et al., 2011). With the sequencing of additional PRIs, more classes of accessory genes have been observed, although they tend to fit within four main functional groups: transcriptional regulation, electron transport, oxidative stress resistance, and molybdenum cofactor biogenesis (Figure 5).

Transcriptional regulation in response to environmental was identified as a possible role for genes in the PRI, specifically via a histidine kinase signaling pathway (Melnyk et al., 2011). This was confirmed using genetics, as *pcrPSR* are each independently required for perchlorate reduction in *Azospira suillum* PS (Melnyk et al., 2013). The *pcrPSR* genes in *Propionivibrio militaris* MP, *Dechloromonas aromatica* RCB, and *Azospira* spp. form a monophyletic group with the *pcrSR* genes from *Magnetospirillum* spp. PRIs. The orientation of *pcrS* and *pcrR* on different coding strands is conserved in this group, but *Sedimenticola* spp. and *Dechlorobacter hydrogenophilus* LT-1

each have histidine kinase systems with unique genomic and domain organization (Figure 5). Nevertheless, we have used the *pcrPSR* nomenclature developed in PS for these organisms to illustrate that they may play a similar role in transcriptional regulation. In the Alphaproteobacteria PRB, there is the possibility of other transcriptional regulators involved with perchlorate reduction as *Azospirillum* sp. TTI contains a *crp* homolog and the *Magnetospirillum* spp. contain a *lysR* homolog (Figure 5).

The exact mechanism of electron transport from the inner membrane to the periplasmic perchlorate reductase is unknown. PcrC is a soluble tetraheme cytochrome previously thought to be an important part of perchlorate reductase function (Bender et al., 2005). The *pcrC* gene is essential for perchlorate reduction in *Azospira suillum* PS (Melnyk et al., 2013), although its complete absence in the genome of *Arcobacter* sp. CAB indicates that it is not universally required for perchlorate reductase function (Carlström et al., 2013). The electron transport mechanism is difficult to infer from genomic observation alone, as PRIs have many genes that could be involved in transporting reducing equivalents (red genes, Figure 5). Two genes occur in multiple PRIs: *pcrQ,* which encodes a quinol dehydrogenase tetraheme cytochrome similar to well-characterized proteins such as CymA, NrfH, and NapC (Simon and Kern, 2008), and *pcrO,* which is homologous to the gamma subunit of ethylbenzene dehydrogenase (Kloer et al., 2006). Both of these genes have slight defects when grown on rich media containing perchlorate (Melnyk et al., 2013), which are more pronounced on minimal media.

Oxidative stress defence has recently emerged as an important fitness determinant during perchlorate reduction (Chapter 3), likely due to the inadvertent accumulation of the side product hypochlorite from chlorite dismutation (Hofbauer et al., 2014). The key genes in this functional group are the *sigF/nrsF* sigma factor/anti-sigma factor system, the putative methionine sulfoxide reductase genes *yedYZ*, and the methionine-rich peptide gene *mrpX,* which have been proposed to form a periplasmic methionine-cycling mechanism to detoxify hypochlorite (Chapter 3). New PRIs have reiterated the importance of this function, as MP and LT-1 contain homologs of the methionine sulfoxide reductase genes *msrA* and *msrB* (Figure 5). However, unlike the cytoplasmic housekeeping version of these genes, they are predicted to be exported to the periplasm, where the bulk of perchlorate-associated oxidative stress is likely to occur.

Perchlorate reductase is part of the diverse DMSO reductase family and thus contains a molybdenum ion bound by two pterin guanine dinucleotide groups, also known as the molybdenum cofactor, or Moco (Rothery et al., 2008; Schwarz et al., 2009). Within the PRI of every PRB in the Alpha- and Betaproteobacteria, there is an *moaA* homolog (Figure 5). MoaA is a radical SAM-dependent enzyme which catalyzes the first step in pterin synthesis from GTP (Hänzelmann and Schindelin, 2006; Schwarz et

al., 2009). All of the known PRB isolates contain a homolog of *moaA* elsewhere in the genome, however, the PRI *moaA* gene appears to be important for perchlorate reduction in minimal media at least in the case of *Azospira suillum* PS (Melnyk et al., 2013).

PRIs are incorporated at distinct genomic sites, even within a family

The exact boundaries of the PRI outside of the core were not delineated in its initial observation, but genes associated with mobilization were found outside of the core (Melnyk et al., 2011). Variability in mobilization genes was even observed between PRIs with similar cores, suggesting that the core genes can be found in various types of mobile elements (Melnyk et al., 2011). In the first complete genome, *Azospira suillum* PS, the PRI core genes are located close to a proline tRNA gene (Melnyk et al., 2011). Upon closer inspection, this gene includes a 48-bp sequence which is also found at the other end of the PRI (Figure 6). Using this region as a scaffold, we were able to assemble a contig with the same genomic organization in the other two *Azospira* PRB strains, KJ and ZAP. We noticed that the genes outside of these repeats were orthologs detected in all Rhodocyclaceae (e.g. *ispB*, *pilSR*) so we looked at their organization in other strains. These flanking genes frequently form a contiguous region in both other PRB and non-perchlorate reducing strains such as *Azospira oryzae* 6a3, confirming that the PRI was integrated into the proline tRNA (Figure 6). In other Rhodocyclaceae, however, this operon is not intact but has had a different genomic island incorporated, suggesting that this specific tRNA is a frequent target of mobile elements. Outside of the PRI core but within the boundary formed by the direct repeats is a site-specific recombinase homologous to *xerD* (Carnoy and Roten, 2009; Domínguez et al., 2011), thus providing a possible mechanism for PRI integration.

The boundaries of the entire PRI were unclear in the complete genome of *Dechloromonas aromatica* RCB, which lacked homologs of the non-core genes from the *Azospira* spp. PRI. However, using our ortholog detection approach to compare the genome of RCB to its close relative, *Dechloromonas* sp. JJ, we were able to delineate the exact boundaries of the PRI in RCB (Figure 7). The RCB PRI is much larger, spanning 145 kB and containing genes for plasmid replication and conjugation. This structure denotes an integrative and conjugative element, or ICE, and may be capable of catalyzing its own excision, transmission and perhaps replication (Burrus et al., 2002).

Due to the fragmentary nature of draft genome assemblies, we were unable to confidently map the location of the PRI insertion in *Propionivibrio militaris* MP or *Dechloromonas agitata* CKB. In *Dechlorobacter hydrogenophilus* LT-1, the PRI is inserted in a region with conserved synteny in other Rhodocyclaceae (Figure 8). However, there are no obvious 'scars' or evidence for a mechanism of integration and no close non-perchlorate reducing relative to LT-1, so the history of its PRI is unclear. The CKB PRI is located on its own relatively small contig, but re-mapping of paired-end reads to the

contigs using bowtie indicates that it is flanked by near-identical insertion sequences, thus forming a composite transposon. This organization was seen in chlorate-reducing bacteria, where the genetic 'cargo' for the metabolism is located between two insertion sequences, which enable its mobilization (Clark et al., 2013).

The PRIs from several isolates within the genera* Azospira *and* Magnetospirillum *are the result of recent HGT

The conservation of the PRI in three of the four sequenced *Azospira* species and its integration into an identical location in all three (Figure 6), is suggestive of one acquisition event prior to the divergence of the four strains, followed by loss of the PRI in strain 6a3. However, the nucleotide comparisons between the *Azospira* strains do not support such a history. By mapping SNPs in the three *Azospira* draft sequences (KJ, ZAP, and 6a3) to the reference PS genome, we see that the PRI in KJ and ZAP has very few SNPs, especially when compared to the conserved genes outside the PRI, which have mutation rates roughly average for what is expected from the genome as a whole (Figure 6). The lack of nucleotide divergence between the three PRIs, even at synonymous sites and intergenic regions indicates that the PRI was acquired independently and recently by the three extant *Azospira* PRB isolates. Of note is that two of the six SNPs in the PRI result in amino acid substitutions in the sequence of PcrA, the initial reductase in the perchlorate reduction pathway (I431L in *Azospira* sp. ZAP, Y736H in *Azospira* sp. KJ), raising the possibility that parts of the PRI may still be evolving under positive selection.

A similar situation exists between *Magnetospirillum bellicus* VDY and *Magnetospirillum* sp. WD, however, in this case it appears that the PRI core has recombined between two unique genomic islands (Figure 9). In the VDY PRI, the core genes are flanked by 300 base pairs of homology (Figure 9, red hash lines), which could indicate where the PRI core was integrated into a previously existing genomic island. However, the WD PRI has only one SNP across a nearly 20 kB region shared with VDY (blue dashed lines, Figure 9). This region extends beyond the PRI core and beyond the putative core boundaries of VDY, including part of a large hypothetical gene (Figure 9). This indicates that the recombination between the VDY and WD islands occurred after the acquisition of the PRI core by the VDY island. Again, the near-perfect nucleotide identity of this region indicates that despite VDY and WD's close phylogenomic relationship (Figure 2), the PRI core has been acquired recently via horizontal gene transfer.

PcrAB and Cld: Monophyletic anchors of an ancestral PRI

Although the *pcrAB* and *cld* genes found in perchlorate-reducing bacteria have phylogenies suggestive of horizontal gene transfer within the hisory, they both are

largely monophyletic and thus support a single ancestral PRI (Figures 10 and 11). The only exception to this is the co-option of the *cld* gene sequence in the formation of chlorate-reduction transposons (e.g in *Ideonella dechloratans* and *Alicycliphilus denitrificans*); however, it seems likely that these came from We rooted the Cld sequences from PRB and two chlorate-reducing bacteria by using a metagenomic sequence and the Cld sequence from '*Candidatus* Nitrospira defluvii' (Maixner et al., 2008; Kostan et al., 2010). While these outgroup sequences likely encode periplasmic proteins, they are not linked with *pcrAB* genes genomically, nor are they found in known perchlorate-reducing bacteria. A more comprehensive phylogeny of Cld sequences indicates that these sequences are intermediates between PRB Clds and the broadly distributed, but generally cytoplasmic Clds (Mlynek et al., 2011; Clark et al., 2013). The PcrAB sequences were concatenated and used to construct an alignment rooted to three sequences of Archaeal periplasmic nitrate reductases (Yoshimatsu et al., 2000; 2002; Martinez-Espinosa et al., 2007), which are closely related to extant PcrAB sequences (Clark et al., 2013).

Convergent evolution has led to polyphyletic accessory genes in the PRI

While PcrAB and Cld are found in every PRI, there are many genes that are not found in every PRI but are present in more than one (Figure 5). We were interested in attempting to map acquisition of these genes on to the evolutionary history, that is, attempting to understand whether these genes represented evolutionary innovations that were then propagated horizontally. Although we only focus on two specific examples here for which we have functional data, we found that these genes are largely polyphyletic and likely represent dozens of independent acquisitions. Frequently the origin of these accessory genes is from the 'host' chromosome, providing indirect support for the argument that HGT of the PRI within a narrow taxonomic scope is frequent.

Co-opting genes from denitrification to build a new electron transport pathway

In the PRI cores of all sequenced PRB, with the exception of *Arcobacter* sp. CAB, there is a gene predicted to encode a tetraheme *c*-type cytochrome of the NapC/NirT family. These proteins are integral membrane proteins, which perform an important function in diverse electron transport chains: providing a pathway for reducing equivalents to move between the membrane-bound quinone pool and periplasmic oxidoreductases (C. Myers and J. Myers, 1997; Simon and Kern, 2008; Kim et al., 2008). The presence of this gene in most PRIs underscores the adaptive benefit that a dedicated electron transport chain provides to perchlorate reducers. Despite this conserved role, a phylogeny of the PcrQ sequence and other close relatives indicates

that it is polyphyletic despite its conserved synteny, indicating multiple acquisitions of PcrQ in the PRI evolutionary history (Figure 12).

Many of the sequences closely related to PcrQ are not from perchlorate reducers, thus their function is presumably different from that of PcrQ. A survey of the genomic location of these *pcrQ* homologs indicates that the majority are found in two conserved genomic clusters. The first is the *nir* cluster that has been described in *Pseudomonas stutzeri* (Figure 12, PSJM300_03235) (Jüngst et al., 1991). A variation of the *nir* cluster is seen in all four of the perchlorate-reducing strains and may be the ancestral donor of the *pcrQ* gene to the Alpha- and Gammaproteobacteria (Figure 12).

Certain homologs of *pcrQ* can be found adjacent to a cytochrome *c* biogenesis operon (*ccmABCDEFGHI*) in PRB or their close relatives (Figure 12, NapQ sequences). This cluster exists in *Azospira suillum* PS and previous TN-seq fitness experiments showed that these genes were important for growth on nitrate as the sole electron acceptor (unpublished data). A survey of the canonical denitrification pathways in PS fails to find a *nar* cluster and reveals a *nap* operon organization of *napDAGHB* (Dsui_507-Dsui_511) in which *napC* is notably absent. This organization is also seen in *Shewanella oneidensis* MR-1, where the CymA tetraheme cytochrome has taken over the NapC role of quinol dehydrogenase (Schwalb et al., 2003; Gao et al., 2009). Interestingly, a deletion of *pcrQ* is only slightly impaired for growth on perchlorate, and is similar to wild-type when grown diauxically with nitrate and perchlorate (Melnyk et al., 2013). We proposed that the presence of nitrate may lead to cross-regulation and complementation of a *pcrQ* defect by expression of this alternative quinol dehydrogenase, which we have designated NapQ.

To test this hypothesis, we generated single mutants of *pcrQ* and *napQ*, as well as double deletion of both (Δ*pcrQ*Δ*napQ*). We grew these strains anaerobically on rich ALP media containing both nitrate and perchlorate. We found that the Δ*pcrQ* and Δ*napQ* single deletions lagged longer that wild-type PS, but were able to reach a higher final yield (Figure 13). The difference in final yield may be due to slower but more efficient use of electron acceptor, which is the limiting factor under these conditions. However, the Δ*pcrQ*Δ*napQ* double deletion was not able to grow at all under these conditions (Figure 13), suggesting that the recently duplicated *napQ* and *pcrQ* genes have partially redundant functions despite being co-opted by different electron transport pathways.

The putative methionine sulfoxide reductase YedY is sourced from two anciently diverging clades

The *yedY* gene from the *Azospira suillum* PS PRI is an important determinant of reactive chlorine species and has been proposed to be a periplasmic methionine sulfoxide reductase (Chapter 3). However, it is frequently found outside of the PRI in

PRB as well as other bacteria, and a phylogenetic analysis suggests that YedY sequences can be partitioned into two anciently diverging groups: one in which it serves a 'housekeeping' role in maintaining periplasmic methionine redox homeostasis, and one where it is frequently found in mobile elements associated with specific bacterial lifestyles, such as the PRI and certain pathogenicity islands (Chapter 3). A survey of the YedY sequences found in PRIs shows that there are representatives from both groups, indicating polyphyly on a much more ancient scale than that exhibited in PcrQ. The YedY sequences from the Betaproteobacteria appear to be derived from other Rhodocyclaceae (Figure 14). Similarly, the YedY sequence found in the *Magnetospirillum* PRI appears to be derived from other *Magnetospirillum* isolates, both perchlorate-reducing and not (Figure 15). The non-PRI YedY sequences mimic the divergence pattern seen in the whole-genome MLSA (Figure 2), suggesting that the extant *Magnetospirillum* PRI originated after the formation of the *Magnetospirillum* genus, but prior to the divergence of strains WD and VDY. Despite being also in the Alphaproteobacteria, the YedY sequences from *Azospirillum* sp. TTI are very similar to those from other *Azospirillum* species (data not shown), indicating that Alphaproteobacteria PRIs may be genus-specific.

Discussion

This study represents a major increase in the number of sequenced genomes of perchlorate reducers and signifies the first comprehensive analysis of both mechanisms of HGT associated with this metabolism and the origins of accessory genes. As such, it has provided new insights into the evolution and ecology of perchlorate reducers and has allowed the generation of hypotheses concerning the origin and propagation of this metabolism.

Why only certain clades?

Although perchlorate reduction is distributed broadly among four different classes within the Proteobacteria, the taxonomic scope of isolates within those classes is narrow (Figures 2-4). The reason for this is unknown, but we suspect several factors are at play. First, it is very likely that isolation and enrichment techniques are biased for a certain type of PRB, likely a copiotroph that has a fast growth rate. It may be that these organisms are only one of several microbial strategies that make use of perchlorate reduction. Indeed, it has previously been observed that altering traditional enrichment parameters such as pH, temperature, or salinity is sufficient to alter the perchlorate reducers isolated (Thrash et al., 2010; Carlström et al., 2013). Additionally, 16S or metagenomics-based analyses of perchlorate-reducing communities may uncover a taxonomically richer set of organisms than those currently present in PRB isolate collections; this avenue has recently been taken by several research groups studying

perchlorate-removing bioreactor communitites (Van Ginkel et al., 2010; Xiao et al., 2010).

An organism may also require additional traits in order to become a successful perchlorate-reducer. Generally speaking, perchlorate reducers are motile aquatic organisms, can utilize a variety of electron donors (and acceptors), and are generally not fermentative (Coates and Achenbach, 2004). It may be that only certain taxonomic groups such as Rhodocyclaceae and Rhodospirillaceae are prepared to be successful in environments such as these, and thus these groups harbor the most cosmopolitan PRB.

Another possibility is that this taxonomic restriction is enforced by the mechanism of PRI integration. The *Azospira* PRI is integrated at a very specific site in a tRNA, and the exact sequence of the site is not conserved in other Rhodocyclaceae. This specificity may mean that *Azospira* has an advantage in certain environments, where it can acquire a version of the PRI easily through site-specific integration at a neutral allele. Furthermore, our phylogenetic analysis indicates that accessory genes from the PRI are often related to genes found in the host organism. We hypothesize that this streamlines PRI integration with host metabolism, specifically in the realms of regulation and electron transport. Specificity of both integration mechanisms and accessory genes results in a positive feedback loop where a PRI becomes more and more optimized for a given lineage. Ecologically, this may have the effect of reducing PRI diversity, as an *Azospira* strain that has acquired a PRI tailored for its needs will be much more competitive than an organism attempting to piece together a PRI *de novo*.

Why is perchlorate reduction seemingly always the result of horizontal gene transfer?

Despite the tailored nature of the PRI in *Azospira*, the lack of the island in strain 6a3 and the obvious horizontal transfer of extant PRIs indicates that perchlorate reduction is not a conserved function of this genus, nor is it for any taxonomic group we have studied (Figures 1-5). Additionally, the ubiquitous presence of HGT "scars" associated with perchlorate reduction indicate that it is frequently, and perhaps exclusively propagated via horizontal gene transfer. This stands in stark contrast to reductive pathways such as denitrification, which are present in many different Proteobacteria and vary little in operon organization across large phylogenetic groups, denoting a much greater role for vertical inheritance.

The evolutionary trajectories of these two metabolisms are likely shaped by the abundance of these electron acceptors in the environment. Denitrification is ubiquitous; nitrate is found in many types of aquatic and terrestrial environments and is continuously generated from inorganic nitrogen sources via nitrifying bacteria. Conversely, perchlorate is rarely found in ecologically relevant quantities due to its slow abiotic generation in the atmosphere. The discrepancy between the abundance of

these two compounds mimics the discrepancy in the frequency of the respective pathways in isolated bacteria.

This "rare perchlorate" hypothesis implicitly states that maintaining the PRI in the absence of perchlorate has a fitness cost. The size of the PRI relative to the genome is the most obvious fitness defect associated with maintaining this metabolism in the absence of perchlorate. The *Azospira* genome is approximately 3.5 Mb and the PRI constitutes a 22kb region, which represents a non-negligible fitness cost associated with chromosomal replication during growth. There may also be a fitness cost beyond genome replication incurred by baseline expression of genes, which could lead to inefficient use of amino acids and expensive cofactors (e.g. Fe-S clusters, hemes) or antagonism with existing pathways (e.g. denitrification), as suggested by the genetics of *pcrQ* and *napQ* in *Azospira suillum* PS (Figures 12 and 13). Current work in our lab is underway to understand exactly how perchlorate and nitrate are regulated in PS and if there are further interactions between these metabolisms.

Although perchlorate is much rarer than nitrate in the environment, there are clearly situations where it is adaptive to have the ability to reduce perchlorate. Perchlorate accumulates in dry areas such as the Atacama Desert and the Antarctic Dry Valleys (Kounaves et al., 2010), but also is deposited across much wider geographic scales (Rajagopalan et al., 2009). However, because its concentration in groundwater is tied to stochastic rainfall and irrigation events, many geographical areas with high washout may never reach a biologically relevant concentration of perchlorate, particularly prior to industrial contamination of perchlorate. This results in non-overlapping niches, that is, a microenvironment where PRB are successful may be geographically isolated from other similar microenvironments, thus limiting dispersion of PRB. In other words, we hypothesize that the frequency of environments where PRB are successful is low enough or that selection against the PRI in non-perchlorate environments is negative enough, that the PRI is never "fixated" into a taxonomic group.

The stochastic pattern of natural perchlorate deposition and lack of continuous positive selection provides a plausible hypothesis for why an ancient metabolism would exist in such a low equilibrium abundance among bacterial genomes, yet continue to persist. However, release of anthropogenic perchlorate into the environment may be changing this equilibrium, providing a continuously selective environment for PRB or increasing the abundance of PRIs in the global pangenome. The minute number of SNPs among PRIs from the *Magnetospirillum* and *Azospira* species relative to the rest of the genome indicate a very recent acquisition of the PRI, especially relative to the divergence of the genomes as a whole. One possibility is that these heavily contaminated sites are acting as incubators of PRIs, but not necessarily PRB. While PRB may not be competent for dispersion from site to site, dispersion of the PRI into the

pangenome via phage transduction may be the mechanism driving the HGT of specific PRIs within a genus observed in this study. Further sequencing of genomes and PRIs of isolates and communities from contaminated sites, bioreactors, and pristine environments will be essential in gaining a more complete understanding of the dynamics of PRI evolution.

Ancient evolution of perchlorate reduction

All extant sequences of PcrA and Cld are monophyletic and are linked together on the same "piece" of DNA. This suggests that all modern PRIs descended from a single progenitor PRI. Hypothetically, this progenitor PRI required three critical steps in its assembly. First and most important is the evolution of robust perchlorate reduction activity. Work is currently underway to understand the biochemical nature of the catalytic activity of perchlorate reductase which may provide insight into how it evolved. Once an organism gains the ability to reduce perchlorate, the problem of reactive chlorine species presents itself. The chlorite dismutase enzymatic function likely predates the linkage of *cld* and *pcrA*, as there are Cld sequences basal to the Cld sequences found in PRB. However, Cld sequences from PRB are unique in the sense that they all contain a signal sequence indicating export to the periplasm. We hypothesize that this was another crucial step in the evolution of perchlorate reduction, which allowed chlorite to be dismutated near the site of its production. Once an organism acquired both of these abilites, it became the first perchlorate reducer. However, until these genes becamed genetically linked, this trait was likely inherited only vertically. The first PRI arose when *cld* and *pcrA* were incorporated on the same piece of DNA that could be transferred between unrelated organisms.

We propose that successful events of PRI horizontal gene transfer that lead to propagation within a new taxonomic group are rare (Figure 16). We also propose that less complex PRIs have a much better chance of being incorporated in taxonomically unique backgrounds; the acquisition of lineage-specific accessory genes comes later, once a the PRI is being frequently transferred . An example of the early stages of this progress may be the PRIs in *Sedimenticola* spp. and *Arcobacter* sp. CAB which both contain transposase 'scars' within the PRI core (Figure 5). The similarity in the core proteins PcrAB and Cld in these organisms indicate that their PRI cores diverged relatively recently (Figures 10 and 11).

Recent evolution of perchlorate reduction

Although extant lineages of perchlorate reduction derive from infrequent HGT events, our genomic data indicates that within these lineages recombination and transfer of the PRI is rampant. The most striking instance of this is the near-identity

among the PRIs of the *Azospira* spp. and the exact nucleotide sequence shared by *Magnetospirillum bellicus* VDY and *Magnetospirillum* sp. WD.

The fact that some of the accessory genes are pseudogenes also demonstrates the ongoing evolution of the PRI core (genes with dashed outlines, Figure 5). Inactivation of genes by altering the start codon or introducing a nonsense mutation is an early step in this process, eventually followed by gene loss (Kuo and Ochman, 2010). This likely occurs because genes in the PRI served an adaptive purpose in the donor organism, but are either neutral or deleterious in their current host (Liu et al., 2004).

Our research indicates that PRIs found in a certain taxonomic group tend to source their accessory genes from that same taxonomic group. Indirectly, this provides evidence for frequent HGT within those groups, as it requires many successive transfers and rearrangements to not only link new genes to the PRI, but also generate new operon structures. For example, *pcrQ* has a polyphyletic origin but is located close to the *pcrAB* and *cld* genes, although the exact synteny is not conserved among the different *pcrQ* lineages (Figures 5 and 12). While we now have functional roles for most of the genes in the *Azospira suillum* PS PRI (Chapters 2 and 3), the diversity of PRIs sequenced to date indicates that this metabolism and associated regulatory and stress response mechanisms are evolving idiosyncratically in different lineages.

Conclusion

Perchlorate reduction is a unique metabolism that combines elements of aerobic and anaerobic respiration. The genomic basis for perchlorate reduction, the PRI, is a unique genomic island as well, encoding complex systems for regulation and oxidative stress response. We have shown in this study that the PRI is primarily associated with horizontal gene transfer on varying degrees of time scales. Understanding the dynamics of horizontal transfer of the PRI, as well as how it becomes integrated into host physiology, will be essential for understanding the ecology of this metabolism and where in the environment these genes exist. Towards this goal, we have integrated sequencing of taxonomically diverse isolates and genetic studies of a model system in order to understand the function of the genes in the PRI. Moving forwards, we believe that perchlorate reduction provides an interesting model system for those looking at the intersection of microbial physiology and ecology, and how genomic organization and horizontally transferred genes shape microbial lifestyles.

Materials and methods

MLSA analysis.

Well-known taxa were chosen for illustrative purposes in the Proteobacterial dataset (Figure 1). We eliminated redundancy at the genus level in the Alpha- and Betaproteobacteria datasets prior to enriching the Rhodospirillaceae and Rhodocyclacease families with draft proteomes from the Joint Genome Institute's IMG database (Figures 2-4). These two families are home to all of the perchlorate-reducing isolates in the Alpha- and Betaproteobacteria classes; thus, we hoped that denser taxon sampling of these families would increase our resolution of relationships among perchlorate-reducing bacteria and also provide insight into the origin of accessory genes in the PRI. In each of the four datasets, we included well-established outgroups (e.g. using Firmicutes as outgroups for all Proteobacteria).

To detect orthologs, we used an all-vs-all pairwise BLASTP approach to identify orthologs between each possible pair of proteomes in each dataset encode using inParanoid(Remm et al., 2001). Each pairwise inParanoid output was used as input to the QuickParanoid clustering algorithm to identify clusters of homologous proteins (http://pl.postech.ac.kr/QuickParanoid/). We designated orthologs as those clusters that had exactly one representative from each organism in each dataset (i.e. Proteobacteria, Alphaproteobacteria, Betaproteobacteria, or Rhodocyclaceae). Each set of orthologs was aligned using MUSCLE(Edgar, 2004), and the resulting alignments were processed using Gblocks (Castresana, 2000) running parameters designed to maximize retention of all informative sites (Sassera et al., 2011). The alignments were concatenated and converted to Phylip format and an appropriate model of amino acid substitution was selected using ProtTest 3 (Abascal et al., 2005). Using the best model and the previously selected outgroup, phylogenetic trees were generated using RAxML (Stamatakis, 2014). A tree topology was calculated using a maximum-likelihood method for choosing the best tree from a number of independent inferences. A number of bootstrap replicates were performed as well on each dataset, and these were drawn onto the most likely tree to generate support values for each node. This method is described thoroughly in the RAxML manual. The amino acid subsitution model and the number of independent inferences and bootstraps are given in the figure legends for each of the four datasets.

In order to simplify the viewing of the Alphaproteobacteria and Betaproteobacteria trees, we collapsed taxonomic groups at the order level based on the published taxonomy. In general, there was robust support for all taxonomic orders with the exception of three species from order Rhodobacterales in the Alphaproteobacteria (*Hirschia baltica, Hyphomonas neptunium, and Maricaulis maris*), which claded with the Caulobacterales. This was previously observed for the genome of the species *Hyphomonas neptunium*, which was assigned to the Rhodobacterales based on 16S phylogeny, but groups with the Caulobacterales on the basis of housekeeping gene sequences (Badger et al., 2005). These three species are all members of family

Hyphomonadaceae, which we have depicted in our tree as the order-level classification 'Hyphomonadales' following a previously set precedent (Gruber-Vodicka et al., 2011). We note that this is currently not an official taxonomic grouping, but our analysis suggests that these species should reside in a sister group to the Caulobacterales or perhaps within that order itself.

Custom BioPerl scripts were used to automate many steps of this pipeline, which we plan to make available in the near future (https://github.com/ryanmelnyk). A BioPython script was used to generate gene plots for many of the figures (Clark et al., 2013).

The program Mauve was used to generate whole-genome alignments of the *Azospira* draft genomes (ZAP, KJ, and 6a3) to the finished reference genome of *Azospira suillum* PS, in addition to aligning contigs from the two *Magnetospirillum* draft assemblies (VDY and WD).

Individual gene phylogenies.

For the phylogenies of independent gene first used a HMMER3 protein similarity search to identify similar proteins to the query of interest in our entire HAMAP/IMG dataset (Eddy, 1998). We then took all similar proteins and clustered them with a 90% identity threshold using CD-HIT to reduce the complexity of the dataset (Li and Godzik, 2006), prior to generating an alignment as before, using MUSCLE and Gblocks. We then calculated a best maximum-likelihood tree from 10 independent inferences using RAxML (Stamatakis, 2014). We then viewed the tree using Dendroscope (Huson et al., 2007), and chose an ingroup containing the sequences from the PRI and an appropriate outgroup. The selected sequences were aligned and curated again as before, using MUSCLE and Gblocks. A best amino acid model was determined using ProtTest 3 and a phylogenetic tree with support values was calculated as before using RAxML and the best tree/bootstrap method. In several instances, RogueNaRok (https://github.com/aberer/RogueNaRok) was used to remove rogue taxa that led to ambiguous phylogenetic signals, prior to tree recalculation. None of the "pruned" sequences from the original tree were from PRB or closely related species.

Genetics and growth curves of* pcrQ *and* napQ *deletion mutants.

The Δ*pcrQ Azospira suillum* PS mutant was generated as part of a previous study (Melnyk et al., 2013), whereas the Δ*napQ* and Δ*pcrQ*Δ*napQ* double mutant was generated according to previously published methods (Melnyk et al., 2013). Briefly, this allelic replacement method involves constructing a suicide vector (pNPTS138) with flanking 1 kb regions of the gene to be deleted, followed by electroporation of the vector into PS and subsequent selection of kanamycin-resistant colonies. These colonies are grown up on unselective media, then plated on sucrose-containing media to

counterselect against the presence of the *sacB* sucrose-sensitive allele on the integrated suicide vector. PCR screening is used to distinguish deletion mutants from wild-type revertants.

Growth curves were performed as before (Chapters 2 and 3), but involve diluting back an overnight aerobic culture of PS to an OD_{600} of ~1 prior to a 30-fold dilution into the growth curve media. Generally, 300 μL of this inoculum is placed in a well of a 96-well plate and overlaid with mineral oil to prevent evaporation. The prepared plate is then incubated in a Tecan Sunrise plate reader at 37°C with no shaking in an anaerobic chamber.

References

Abascal, F., Zardoya, R., and Posada, D. (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics (Oxford, England)* **21**: 2104–2105.

Badger, J.H., Eisen, J.A., and Ward, N.L. (2005) Genomic analysis of Hyphomonas neptunium contradicts 16S rRNA gene-based phylogenetic analysis: implications for the taxonomy of the orders "Rhodobacterales" and Caulobacterales. *Int J Syst Evol Microbiol* **55**: 1021–1026.

Bender, K., Shang, C., Chakraborty, R., Belchik, S., Coates, J., and Achenbach, L. (2005) Identification, characterization, and classification of genes encoding perchlorate reductase. *J Bacteriol* **187**: 5090–5096.

Bonaventura, M.P.D., Lee, E.K., deSalle, R., and Planet, P.J. (2010) A whole-genome phylogeny of the family Pasteurellaceae. *Mol. Phylogenet. Evol.* **54**: 950–956.

Burrus, V., Pavlovic, G., Decaris, B., and Guédon, G. (2002) The ICESt1 element of Streptococcus thermophilus belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid* **48**: 77–97.

Carlström, C.I., Wang, O., Melnyk, R.A., Bauer, S., Lee, J., Engelbrektson, A., and Coates, J.D. (2013) Physiological and Genetic Description of Dissimilatory Perchlorate Reduction by the Novel Marine Bacterium Arcobacter sp. Strain CAB. *mBio* **4**:.

Carnoy, C. and Roten, C.-A. (2009) The dif/Xer recombination systems in proteobacteria. *PLoS One* **4**: e6531.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17**: 540–552.

Clark, I.C., Melnyk, R.A., Engelbrektson, A., and Coates, J.D. (2013) Structure and evolution of chlorate reduction composite transposons. *mBio* **4**:.

Coates, J. and Achenbach, L. (2004) Microbial perchlorate reduction: rocket-fueled metabolism. *Nature Reviews Microbiology* **2**: 569–580.

Coates, J.D., Michaelidou, U., Bruce, R.A., O'Connor, S.M., Crespi, J.N., and Achenbach, L.A. (1999) Ubiquity and diversity of dissimilatory (per)chlorate-reducing bacteria. *Appl Environ Microbiol* **65**: 5234–5241.

Domínguez, N.M., Hackett, K.T., and Dillard, J.P. (2011) XerCD-mediated site-specific recombination leads to loss of the 57-kilobase gonococcal genetic island. *J Bacteriol* **193**: 377–388.

Eddy, S. (1998) Profile hidden Markov models. *Bioinformatics (Oxford, England)* **14**: 755–763.

Edgar, R. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* **5**: 113.

Gao, H., Yang, Z.K., Barua, S., Reed, S.B., Romine, M.F., Nealson, K.H., et al. (2009) Reduction of nitrate in Shewanella oneidensis depends on atypical NAP and NRF systems with NapB as a preferred electron transport protein from CymA to NapA. *Isme J* **3**: 966–976.

Gruber-Vodicka, H.R., Dirks, U., Leisch, N., Baranyi, C., Stoecker, K., Bulgheresi, S., et al. (2011) Paracatenula, an ancient symbiosis between thiotrophic Alphaproteobacteria and catenulid flatworms. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 12078–12083.

Hanage, W.P., Fraser, C., and Spratt, B.G. (2005) Fuzzy species among recombinogenic bacteria. *BMC Biol.* **3**: 6.

Hänzelmann, P. and Schindelin, H. (2006) Binding of 5'-GTP to the C-terminal FeS cluster of the radical S-adenosylmethionine enzyme MoaA provides insights into its mechanism. *Proceedings of the National Academy of Sciences* **103**: 6829–6834.

Hofbauer, S., Gruber, C., Pirker, K.F., Sündermann, A., Schaffner, I., Jakopitsch, C., et al. (2014) Transiently produced hypochlorite is responsible for the irreversible inhibition of chlorite dismutase. *Biochemistry*.

Huson, D., Richter, D., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. (2007) Dendroscope: An interactive viewer for large phylogenetic trees. *BMC bioinformatics* **8**: 460.

Jüngst, A., Wakabayashi, S., Matsubara, H., and Zumft, W. (1991) The nirSTBM region coding for cytochrome cd1-dependent nitrite respiration of Pseudomonas stutzeri consists of a cluster of mono-, di-, and tetraheme proteins. *FEBS letters* **279**: 205–209.

Kim, H., Zatsman, A., Upadhyay, A., Whittaker, M., Bergmann, D., Hendrich, M., and Hooper, A. (2008) Membrane tetraheme cytochrome c(m552) of the ammonia-oxidizing nitrosomonas europaea: a ubiquinone reductase. *Biochemistry* **47**: 6539–6551.

Kloer, D.P., Hagel, C., Heider, J., and Schulz, G.E. (2006) Crystal structure of ethylbenzene dehydrogenase from Aromatoleum aromaticum. *Structure* **14**: 1377–1388.

Kostan, J., Sjöblom, B., Maixner, F., Mlynek, G., Furtmüller, P.G., Obinger, C., et al. (2010) Structural and functional characterisation of the chlorite dismutase from the nitrite-oxidizing bacterium "Candidatus Nitrospira defluvii": identification of a catalytically important amino acid residue. *Journal of structural biology* **172**: 331–342.

Kounaves, S.P., Stroble, S.T., Anderson, R.M., Moore, Q., Catling, D.C., Douglas, S., et al. (2010) Discovery of natural perchlorate in the Antarctic Dry Valleys and its global implications. *Environ Sci Technol* **44**: 2360–2364.

Kuo, C.-H. and Ochman, H. (2010) The extinction dynamics of bacterial pseudogenes. *PLoS genetics* **6**:.

Li, W. and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large

sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)* **22**: 1658–1659.

Liu, Y., Harrison, P.M., Kunin, V., and Gerstein, M. (2004) Comprehensive analysis of pseudogenes in prokaryotes: widespread gene decay and failure of putative horizontally transferred genes. *Genome Biol.* **5**: R64.

Maixner, F., Wagner, M., Lücker, S., Pelletier, E., Schmitz-Esser, S., Hace, K., et al. (2008) Environmental genomics reveals a functional chlorite dismutase in the nitrite-oxidizing bacterium 'Candidatus Nitrospira defluvii'. *Environ Microbiol* **10**: 3043–3056.

Martinez-Espinosa, R.M., Dridge, E.J., Bonete, M.J., Butt, J.N., Butler, C.S., Sargent, F., and Richardson, D.J. (2007) Look on the positive side! The orientation, identification and bioenergetics of "Archaeal" membrane-bound nitrate reductases. *FEMS Microbiol Lett* **276**: 129–139.

Melnyk, R.A., Clark, I.C., Liao, A., and Coates, J.D. (2013) Transposon and deletion mutagenesis of genes involved in perchlorate reduction in Azospira suillum PS. *mBio* **5**: e00769–13.

Melnyk, R.A., Engelbrektson, A., Clark, I.C., Carlson, H.K., Byrne-Bailey, K., and Coates, J.D. (2011) Identification of a perchlorate reduction genomic island with novel regulatory and metabolic genes. *Appl Environ Microbiol* **77**: 7401–7404.

Mlynek, G., Sjöblom, B., Kostan, J., Füreder, S., Maixner, F., Gysel, K., et al. (2011) Unexpected Diversity of Chlorite Dismutases: a Catalytically Efficient Dimeric Enzyme from Nitrobacter winogradskyi. *J Bacteriol* **193**: 2408–2417.

Myers, C. and Myers, J. (1997) Cloning and sequence of cymA, a gene encoding a tetraheme cytochrome c required for reduction of iron(III), fumarate, and nitrate by Shewanella putrefaciens MR-1. *J Bacteriol* **179**: 1143–1152.

Rajagopalan, S., Anderson, T., Cox, S., Harvey, G., Cheng, Q., and Jackson, W.A. (2009) Perchlorate in wet deposition across North America. *Environ Sci Technol* **43**: 616–622.

Remm, M., Storm, C.E., and Sonnhammer, E.L. (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *Journal of molecular biology* **314**: 1041–1052.

Retchless, A. and Lawrence, J. (2010) Phylogenetic incongruence arising from fragmented speciation in enteric bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 11453–11458.

Rothery, R.A., Workun, G.J., and Weiner, J.H. (2008) The prokaryotic complex iron-sulfur molybdoenzyme family. *Biochim Biophys Acta* **1778**: 1897–1929.

Sassera, D., Lo, N., Epis, S., D'Auria, G., Montagna, M., Comandatore, F., et al. (2011) Phylogenomic Evidence for the Presence of a Flagellum and cbb3 Oxidase in the Free-Living Mitochondrial Ancestor. *Mol Biol Evol*.

Schwalb, C., Chapman, S.K., and Reid, G.A. (2003) The tetraheme cytochrome CymA is

required for anaerobic respiration with dimethyl sulfoxide and nitrite in Shewanella oneidensis. *Biochemistry* **42**: 9491–9497.

Schwarz, G., Mendel, R.R., and Ribbe, M.W. (2009) Molybdenum cofactors, enzymes and pathways. *Nature* **460**: 839–847.

Simon, J. and Kern, M. (2008) Quinone-reactive proteins devoid of haem b form widespread membrane-bound electron transport modules in bacterial respiration. *Biochemical Society transactions* **36**: 1011–1016.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* **30**: 1312–1313.

Thrash, J.C., Pollock, J., Torok, T., and Coates, J.D. (2010) Description of the novel perchlorate-reducing bacteria Dechlorobacter hydrogenophilus gen. nov., sp. nov.and Propionivibrio militaris, sp. nov. *Appl Microbiol Biotechnol* **86**: 335–343.

Van Ginkel, S.W., Lamendella, R., Kovacik, W.P., Santo Domingo, J.W., and Rittmann, B.E. (2010) Microbial community structure during nitrate and perchlorate reduction in ion-exchange brine using the hydrogen-based membrane biofilm reactor (MBfR). *Bioresource technology* **101**: 3747–3750.

Xiao, Y., Roberts, D.J., Zuo, G., Badruzzaman, M., and Lehman, G.S. (2010) Characterization of microbial populations in pilot-scale fluidized-bed reactors treating perchlorate- and nitrate-laden brine. *Water research* **44**: 4029–4036.

Yoshimatsu, K., Iwasaki, T., and Fujiwara, T. (2002) Sequence and electron paramagnetic resonance analyses of nitrate reductase NarGH from a denitrifying halophilic euryarchaeote Haloarcula marismortui. *FEBS letters* **516**: 145–150.

Yoshimatsu, K., Sakurai, T., and Fujiwara, T. (2000) Purification and characterization of dissimilatory nitrate reductase from a denitrifying halophilic archaeon, Haloarcula marismortui. *FEBS letters* **470**: 216–220.

Tables and Figures

Figure 1. *MLSA tree of Proteobacteria.* A tree depicting the relationships that perchlorate and chlorate-reducing bacteria have with well-studied organisms in the Proteobacteria. The dataset used to generate this phylogeny consisted of 90 taxa and 92 conserved orthologous genes, resulting in an alignment of 13066 amino acid positions. Using the LG substitution model, this tree was determined to be the most likely of 100 independently calculated trees, and all nodes were fully supported over 100 bootstrap replicates.

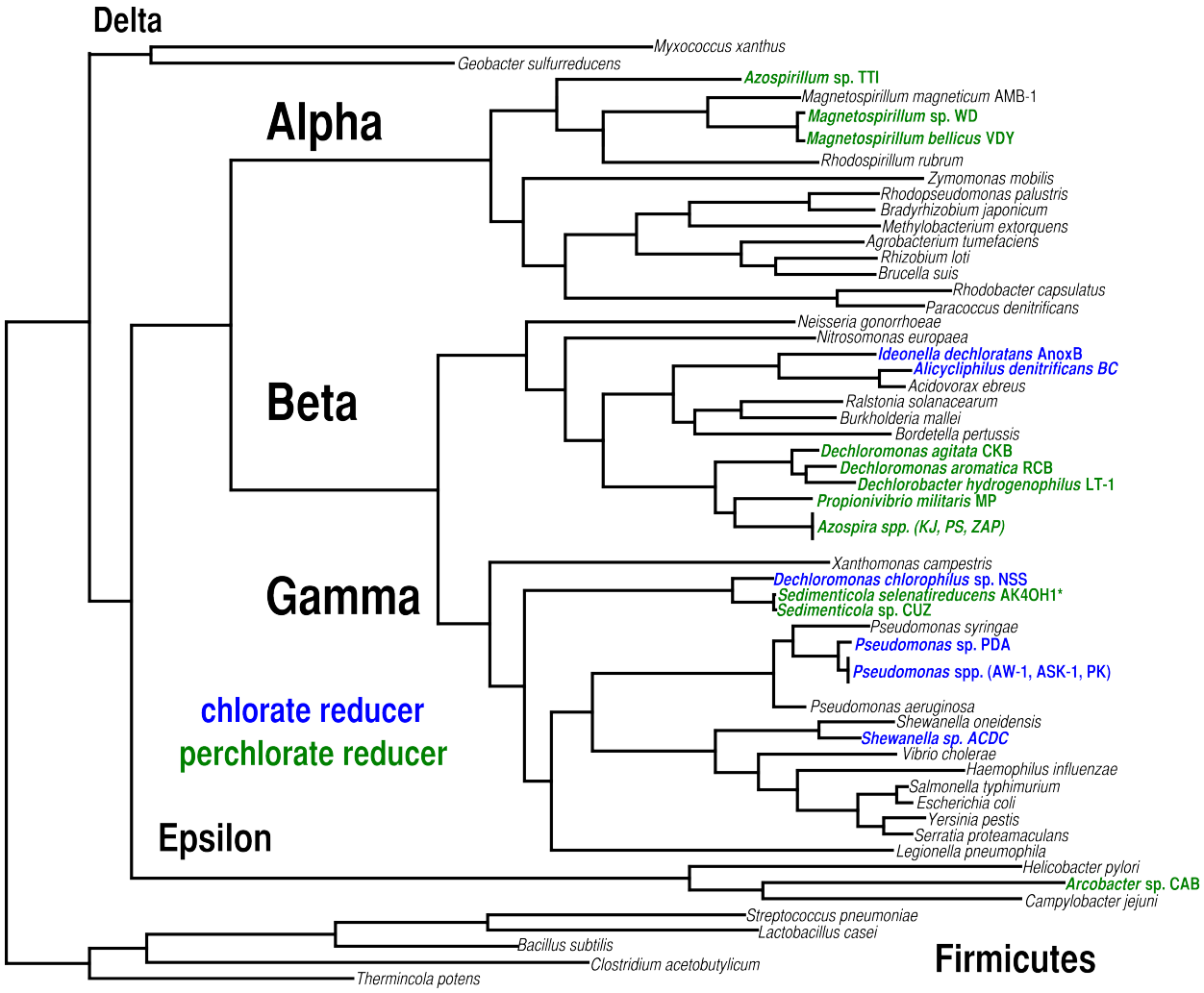


Figure 2. *MLSA tree of the Alphaproteobacteria.* A tree detailing the relationships within the Alphaproteobacteria, focusing on the Rhodospirillales order. The alignment consisted of 9766 amino acids across 80 genes from 96 taxa and the phylogenetic tree was calculated under the LG substitution matrix using 100 bootstraps and 50 inferences to pick the best tree. All nodes had 100% bootstrap support with the exception of the orange node.

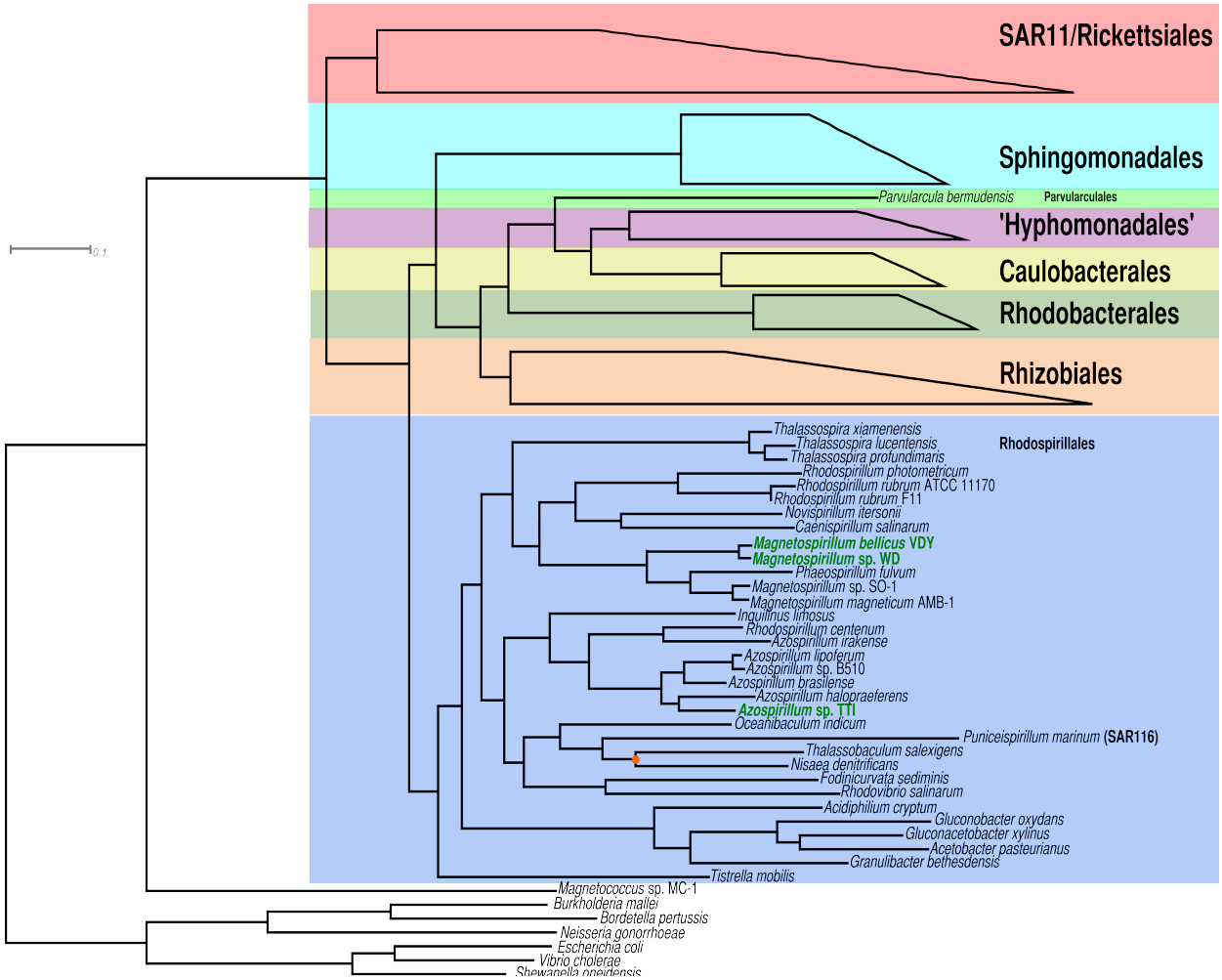


Figure 3. *MLSA tree of the Betaproteobacteria.* A tree depicting the Betaproteobacteria class, with a focus on the Rhodocyclales, the only PRB-containing order. The alignment consisted of 21601 non-redundant amino acid positions from 147 orthologous genes. The tree was calculated using the LG substitution and overlaying the results of 100 bootstrap replicates over the most likely tree of 50 independent inferences. All nodes had 100% bootstrap support.

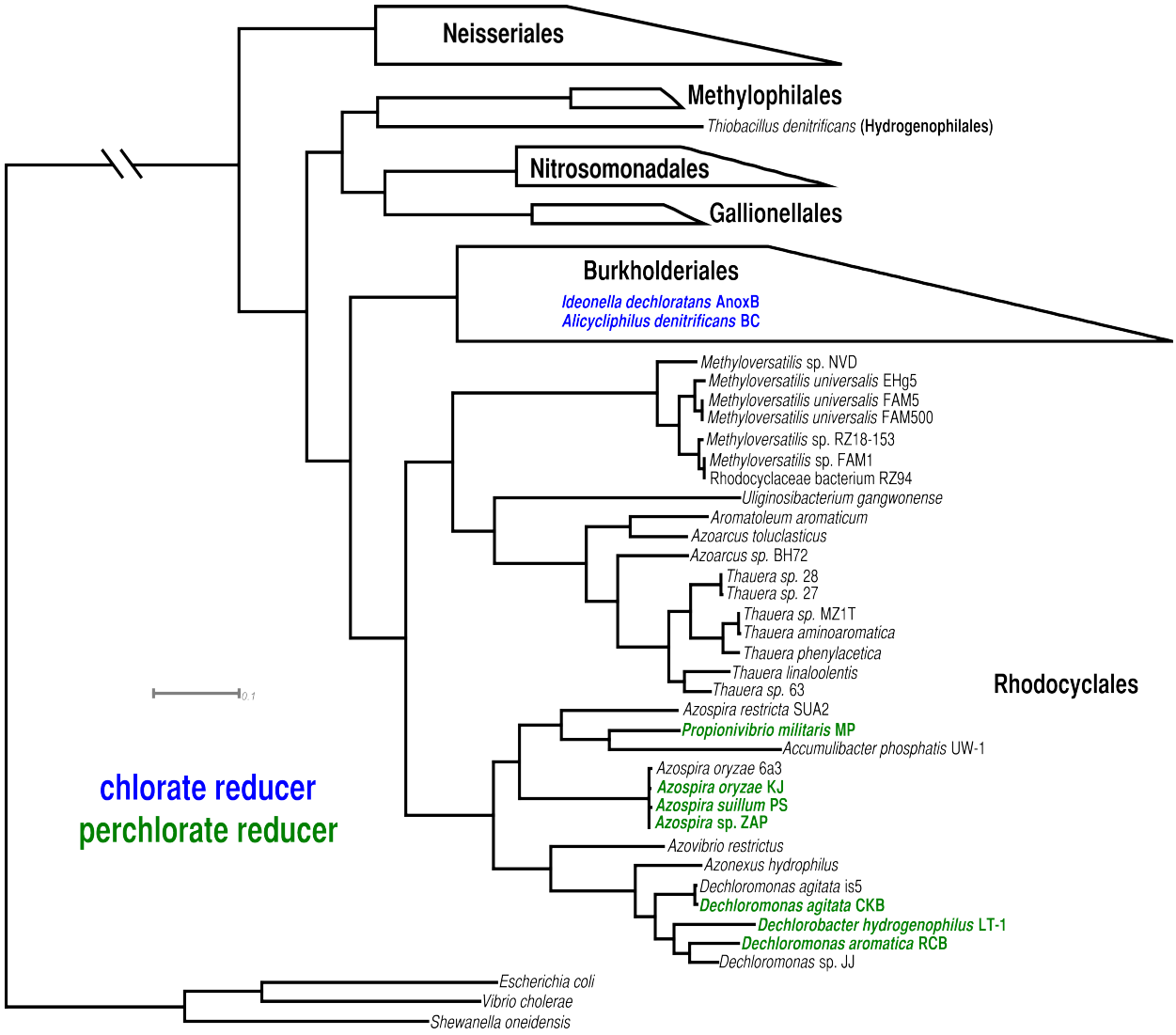


Figure 4. *MLSA tree of selected Rhodocyclales taxa.* An MLSA tree of a clade within the Rhodocyclales containing all PRB. The tree is rooted to taxa from a sister clade within the same order (see Figure 3). The alignment used to construct this tree consisted of 139587 amino acid positions from 229 genes across 19 taxa. The tree was calculated under the LG+F substitution matrix and was generated by drawing 100 bootstrap replicates on the best tree chosen from 50 independent maximum-likelihood inferences. All nodes had 100% bootstrap support.

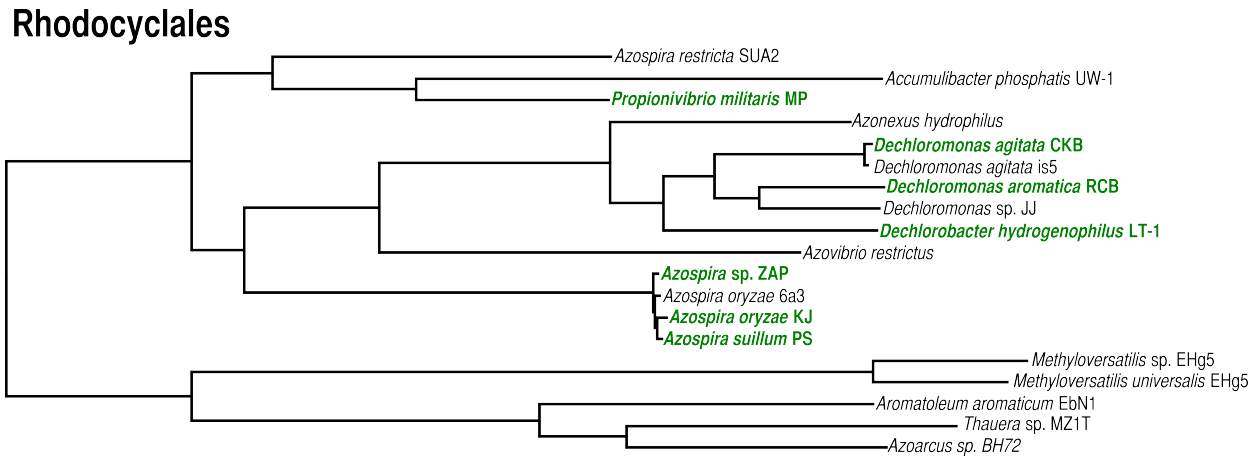


Figure 5. *The cores of the 13 PRIs.* The PRI core from the 13 PRB genomes with genes colored according to functional groups and labeled with gene names, rather than locus tags.

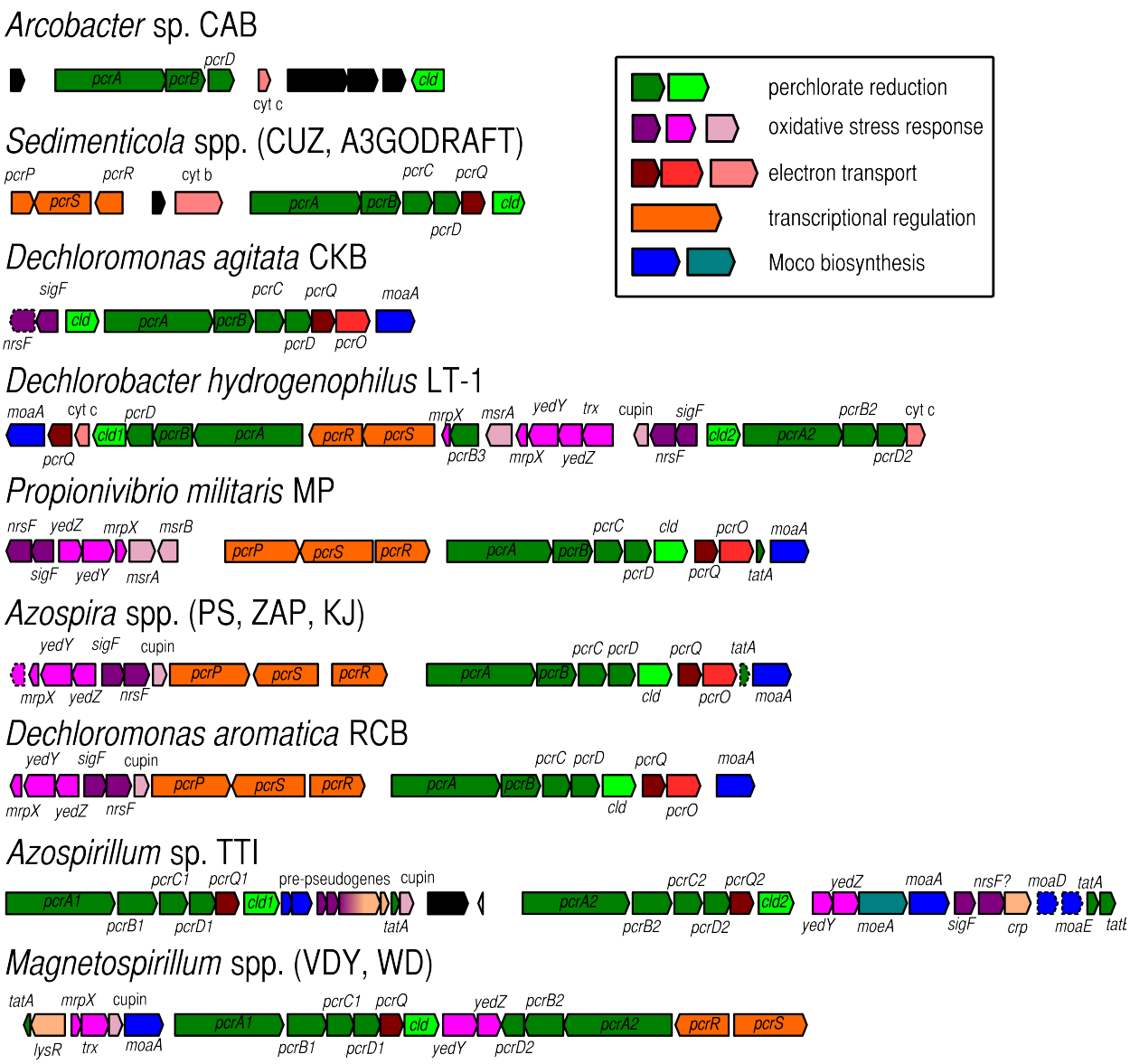


Figure 6. *The PRI structure in* Azospira *spp.* Diagram of the PRI and flanking regions of *Azospira* spp. as well as syntenically conserved genes in the non-PRB *Azospira oryzae* 6a3 and two *Dechloromonas* spp. Overlaid on the PRI is a plot of SNP frequency of three *Azospira* strains relative to the reference genome of *Azospira suillum* PS. Finally, the nucleotide sequence of the putative site of integration within the proline tRNA is shown for all strains (indicated by the red line on the genetic diagrams).

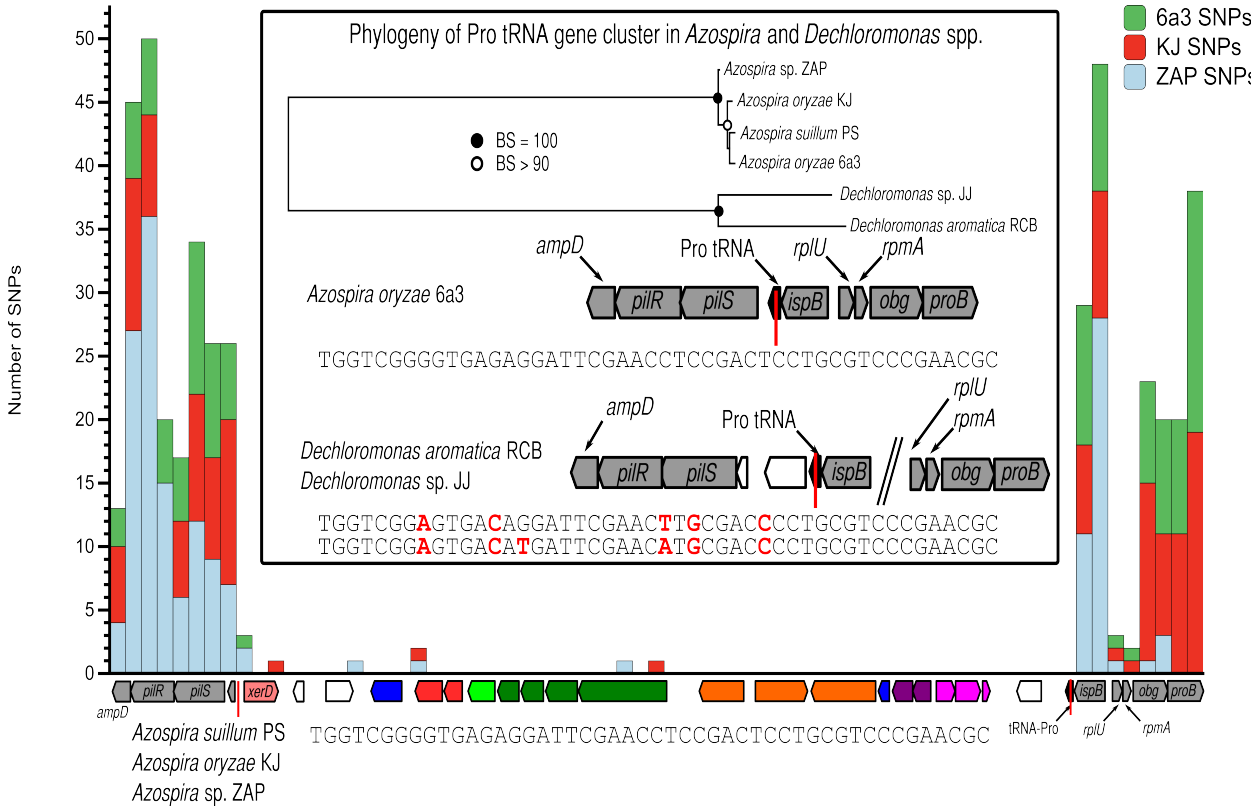
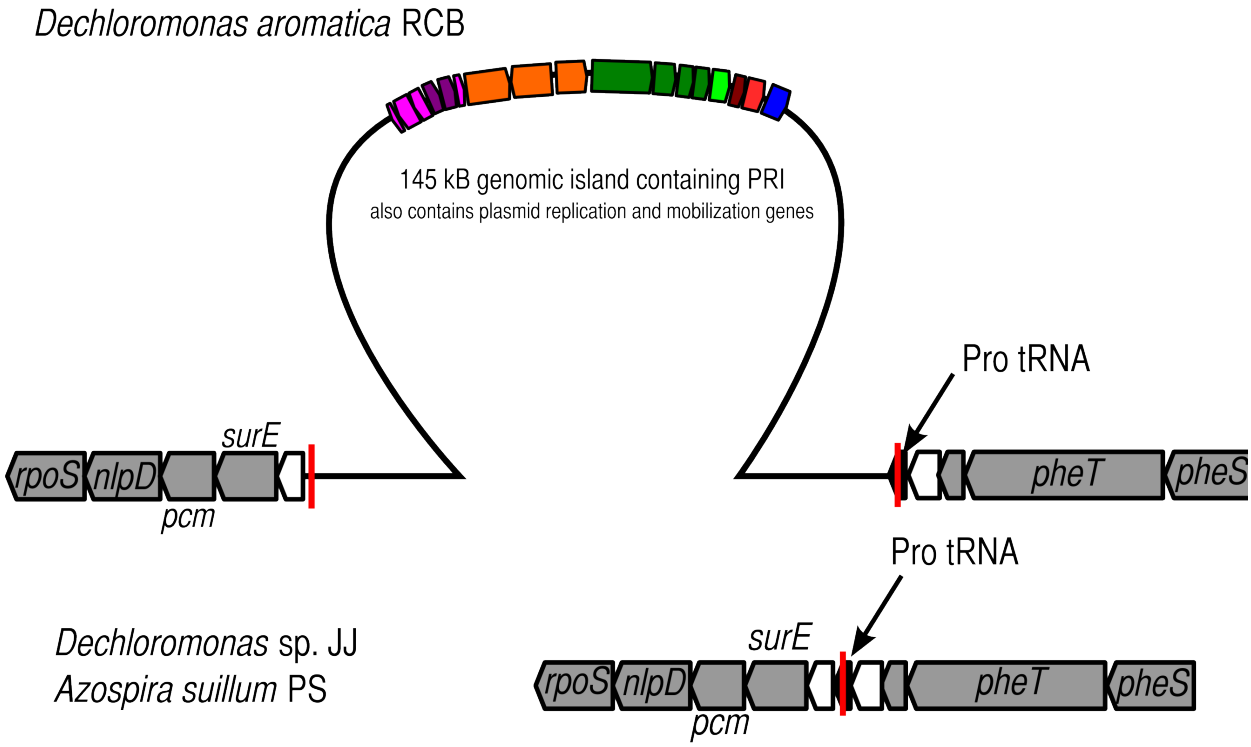


Figure 7. *The conjugative PRI structure of* Dechloromonas aromatica *RCB*. Diagram of the PRI ICE insertion in *Dechloromonas aromatica* RCB in an operon conserved in the non-PRB *Dechloromonas* sp. JJ and *Azospira suillum* PS.



CATGGGGTGCAAGGGGTCGCGAGTTCGAATCCCGCCGCCCCCGACCA

Figure 8. *Structure of the PRI in LT-1.* Diagram of putative PRI insertion in *Dechlorobacter hydrogenophilus* LT-1 and the flanking regions with conserved synteny in other Rhodocyclaceae.

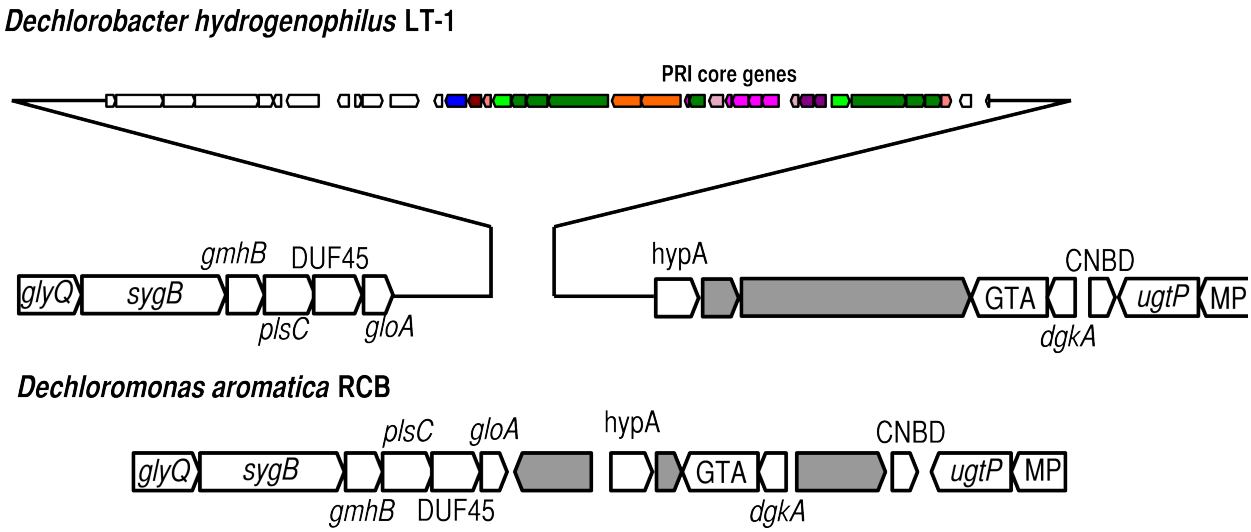


Figure 9. *The conserved core of the* Magnetospirillum *spp. PRIs.* Recombination of core genes between two different *Magnetospirillum* genomic islands to create two separate PRIs.

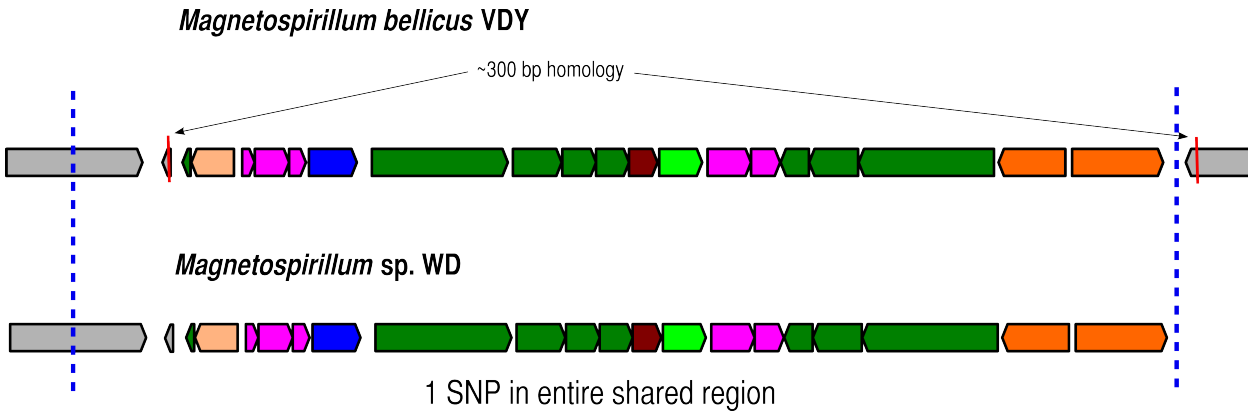


Figure 10. *Phylogenetic tree of Cld sequences.* Phylogenetic tree constructed using an alignment of Cld sequences from PRB and rooted to two sequences predicted to be periplasmic Cld sequences not found in PRB.

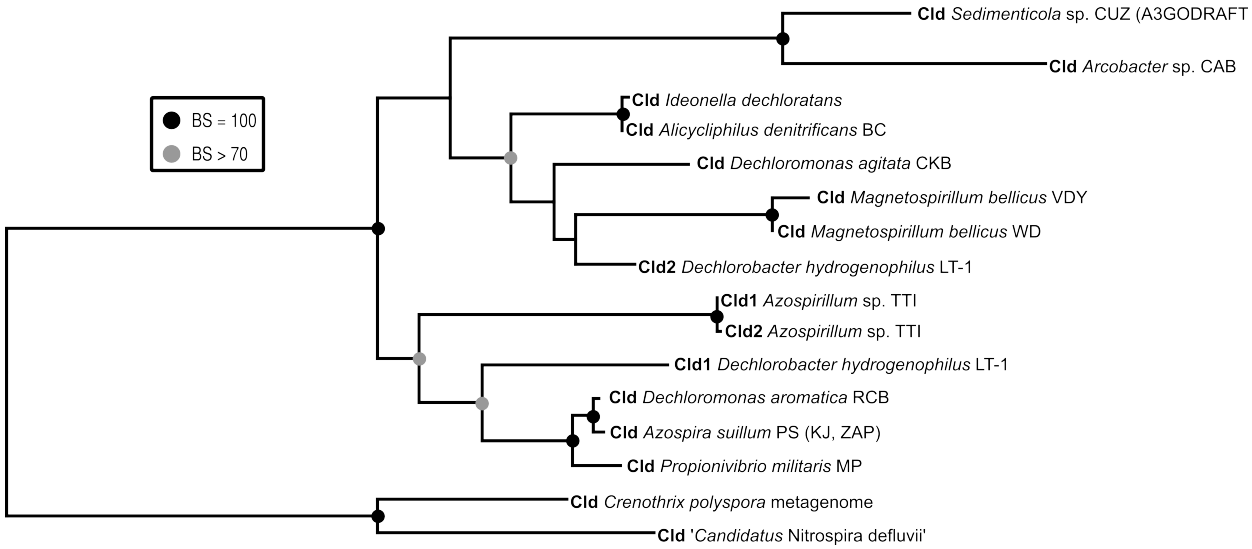


Figure 11. *Phylogenetic tree of PcrAB sequences.* Phylogenetic tree constructed using a concatenation of the PcrA and PcrB sequences from several PRB and rooted to the periplasmic NarGH sequences from several Archaea.

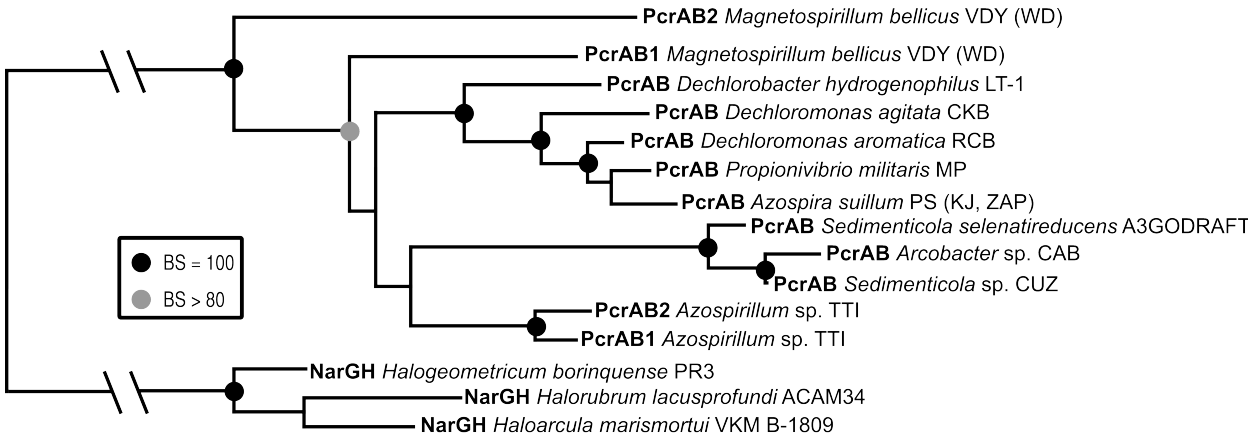


Figure 12. *Phylogenetic tree of PcrQ sequences.* This tree contains the PcrQ sequences from various DPRB rooted to NapC sequences from several Proteobacteria. The boxes at bottom show possible donors of the polyphyletic *pcrQ* gene from VDY and PS.

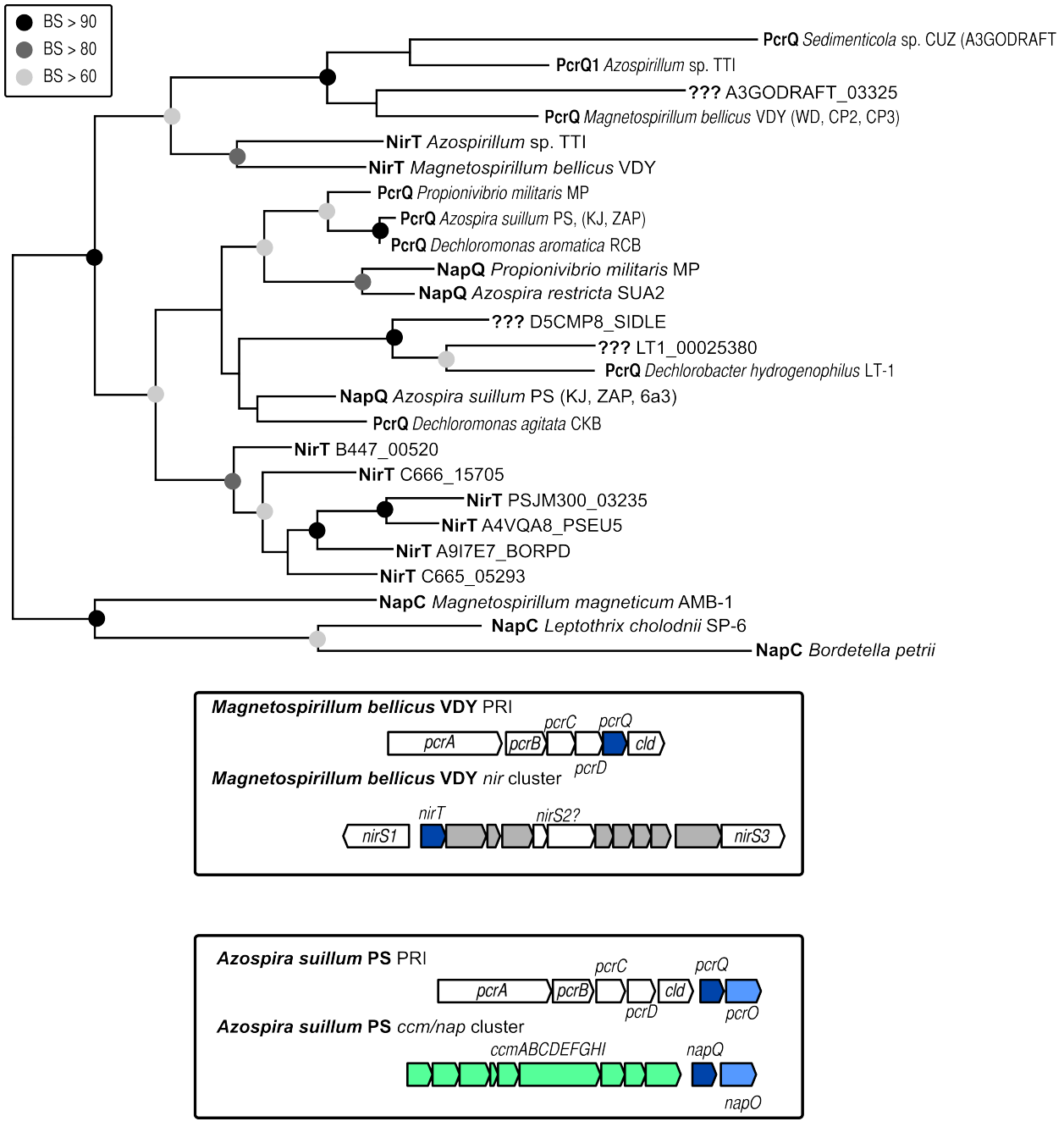


Figure 13. *Possible redundant roles for* napQ *and* pcrQ. Growth curve of wild-type PS compared to Δ*pcrQ*, Δ*napQ*, and Δ*pcrQ*Δ*napQ* strains on rich ALP media containing nitrate and perchlorate.

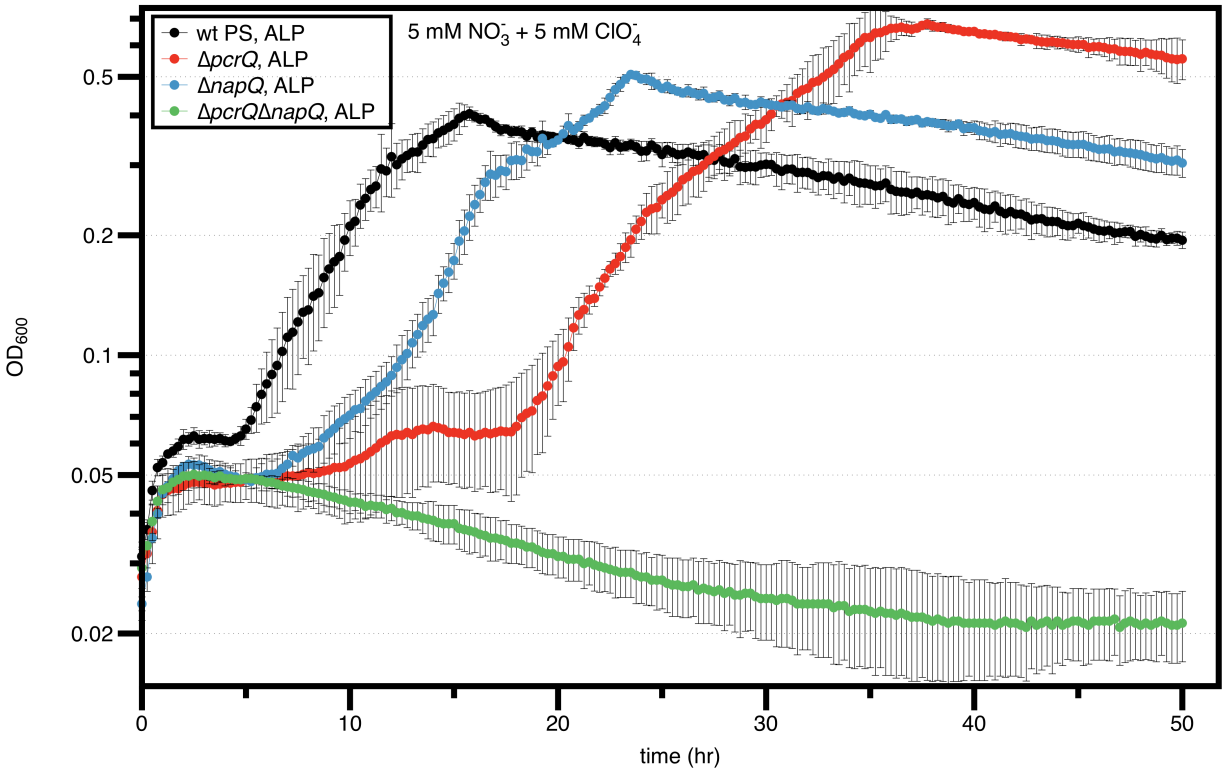


Figure 14. *Phylogeny of YedY from Betaproteobacteria DPRB.* Phylogenetic tree constructed using YedY sequences from Betaproteobacteria DPRB and related sequences.

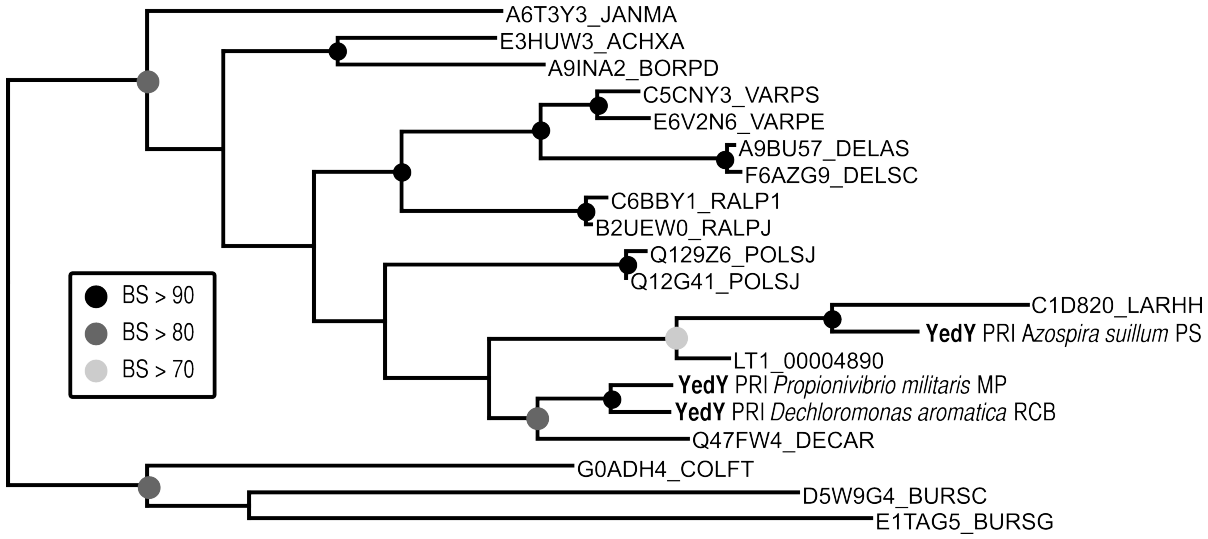


Figure 15. *Phylogeny of YedY from* Magnetospirillum *spp.* Phylogenetic tree constructed using YedY sequences from Magnetospirillum spp. and related sequences.

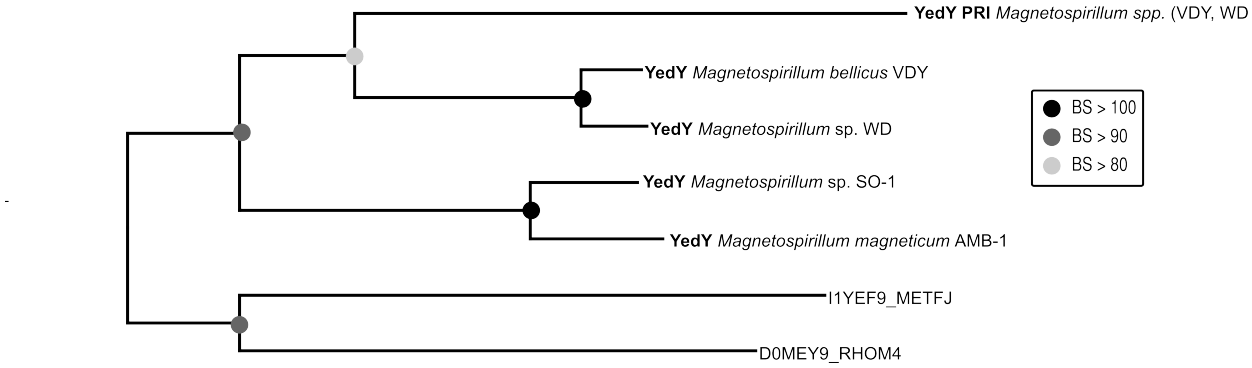


Figure 16. *Conceptual model for the evolution of the PRI.* Rare HGT events of the entire PRI between distantly related organisms likely increase the taxonomic scope of the PRI and perhaps allow perchlorate reduction to occur in new niches (i.e. different substrate utilization, salt water, etc.). However, based on extant genome comparisions, HGT between closely related species is much more common. This frequent but taxonomically restricted HGT optimizes the PRI for function in a particular group, likely by divergent evolution of core proteins like PcrA and Cld, but also by acquisition of lineage-specific accessory genes.

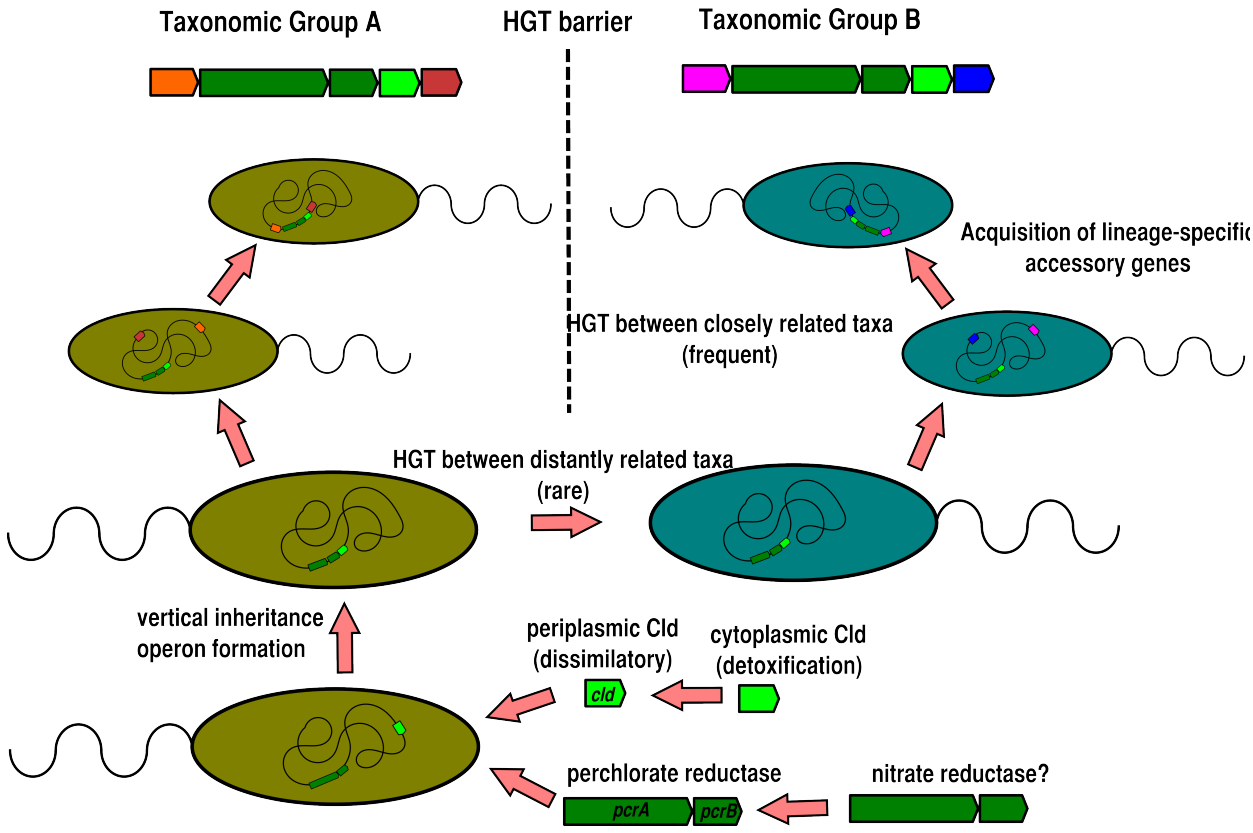


Table 1. *List of locus tags of genes in the PRI.*

gene name	VDY	TTI	RCB	PS	MP	LT-1	CKB	CUZ	CAB
pcrA1	MagVDY_00380	AzoTTI_07766	Daro_2584	Dsui_0149	PropMP_00002000	LT1_00028370	CKB_00037120	CUZ_04206	CAB_00027550
pcrA2	MagVDY_00390	AzoTTI_07782	n/a	n/a	n/a	LT1_00028510	n/a	n/a	n/a
pcrB1	MagVDY_00381	AzoTTI_07767	Daro_2583	Dsui_0148	PropMP_00002010	LT1_00028360	CKB_00037110	CUZ_04205	CAB_00027540
pcrB2	MagVDY_00389	AzoTTI_07783	n/a	n/a	n/a	LT1_00028520	n/a	n/a	n/a
pcrB3	n/a	n/a	n/a	n/a	n/a	LT1_00028410	n/a	n/a	n/a
pcrC1	MagVDY_00382	AzoTTI_07768	Daro_2582	Dsui_0147	PropMP_00002020	n/a	CKB_00037100	CUZ_04204	n/a
pcrC2	n/a	AzoTTI_07784	n/a	n/a	n/a	n/a	n/a	n/a	n/a
pcrD1	MagVDY_00383	AzoTTI_07769	Daro_2581	Dsui_0146	PropMP_00002030	LT1_00028350	CKB_00037090	CUZ_04203	CAB_00027530
pcrD2	MagVDY_00388	AzoTTI_07785	n/a	n/a	n/a	LT1_00028530	n/a	n/a	n/a
cld1	MagVDY_00385	AzoTTI_07771	Daro_2580	Dsui_0145	PropMP_00002040	LT1_00028340	CKB_00037130	CUZ_04201	CAB_00027480
cld2	n/a	AzoTTI_07787	n/a	n/a	n/a	LT1_00028500	n/a	n/a	n/a
pcrP	n/a	n/a	Daro_2587	Dsui_0152	PropMP_00001970	n/a	n/a	CUZ_04211	n/a
pcrS	MagVDY_00392	n/a	Daro_2586	Dsui_0151	PropMP_00001980	LT1_00002839	n/a	CUZ_04210	n/a
pcrR	MagVDY_00391	n/a	Daro_2585	Dsui_0150	PropMP_00001990	LT1_00002838	n/a	CUZ_04209	n/a
moaA	MagVDY_00379	AzoTTI_07791	Daro_2577	Dsui_0141	PropMP_00002080	LT1_00028310	CKB_00037060	n/a	n/a
nrsF	n/a	AzoTTI_	Daro_25	Dsui_01	PropMP_000019	LT1_000	n/a	n/a	n/a

		07793	89	54	00	28480			
sigF	n/a	AzoTTI_07792	Daro_2590	Dsui_0155	PropMP_00001910	LT1_00028490	CKB_00037140	n/a	n/a
yedZ	MagVDY_00387	AzoTTI_07789	Daro_2591	Dsui_0156	PropMP_00001920	LT1_00028450	n/a	n/a	n/a
yedY	MagVDY_00386	AzoTTI_07788	Daro_2592	Dsui_0157	PropMP_00001930	LT1_00028440	n/a	n/a	n/a
mrpX	MagVDY_00376	n/a	Daro_2593	Dsui_0158	PropMP_00001940	LT1_00028430,0002840		n/a	n/a
cupin	MagVDY_00378	AzoTTI_07779	Daro_2588	Dsui_0153	n/a	LT1_00028470	n/a	n/a	n/a
pcrQ1	MagVDY_00384	AzoTTI_07770	Daro_2579	Dsui_0144	PropMP_00002050	LT1_00028320	CKB_00037080	CUZ_04202	n/a
pcrQ2	n/a	AzoTTI_07786	n/a	n/a	n/a	n/a	n/a	n/a	n/a
transposases	n/a	AzoTTI_07780	n/a	n/a	n/a	n/a	n/a	CUZ_04208	CAB_00027490-CAB_00027510, CAB_00027560
msrA	n/a	n/a	n/a	n/a	PropMP_00001950	LT1_00028420	n/a	n/a	n/a
msrB	n/a	n/a	n/a	n/a	PropMP_00001951	n/a	n/a	n/a	n/a
pcrO	n/a	n/a	Daro_2578	Dsui_0143	PropMP_00002060	n/a	CKB_00037070	CUZ_04207	n/a
other cyt c	n/a	n/a	n/a	n/a	n/a	LT1_00028330,0002854	0 n/a	n/a	CAB_00027520
tatA	MagVDY_00374	AzoTTI_07779,7795	n/a	n/a	PropMP_00002070	n/a	n/a	n/a	n/a
tatB	n/a	AzoTTI_07796	n/a	n/a	n/a	n/a	n/a	n/a	n/a
trx	MagVDY_00377	n/a	n/a	n/a	n/a	LT1_00028460	n/a	n/a	n/a

lysR/crp	MagVDY_00375	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
moeA	n/a	AzoTTI_07790	n/a	n/a	n/a	n/a	n/a	n/a	n/a
pre-pseudogenes?	n/a	AzoTTI_07772-07777	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Table 2. *List of locus tags of other genes from the chapter.*

gene name	organism	locus tag
rpoS	*Dechloromonas aromatica* RCB	Daro_2521
nlpD	*Dechloromonas aromatica* RCB	Daro_2522
pcm	*Dechloromonas aromatica* RCB	Daro_2523
surE	*Dechloromonas aromatica* RCB	Daro_2524
hypothetical protein	*Dechloromonas aromatica* RCB	Daro_2525
Pro tRNA	*Dechloromonas aromatica* RCB	Daro_R0067
hypothetical protein	*Dechloromonas aromatica* RCB	Daro_2670
ihfA	*Dechloromonas aromatica* RCB	Daro_2671
pheT	*Dechloromonas aromatica* RCB	Daro_2672
pheS	*Dechloromonas aromatica* RCB	Daro_2673
gloA	*Dechloromonas aromatica* RCB	Daro_3539
sygB	*Dechloromonas aromatica* RCB	Daro_3535
glyQ	*Dechloromonas aromatica* RCB	Daro_3534
gmhB	*Dechloromonas aromatica* RCB	Daro_3536
plsC	*Dechloromonas aromatica* RCB	Daro_3537
DUF45	*Dechloromonas aromatica* RCB	Daro_3538
hypothetical protein A (*hypA*)	*Dechloromonas aromatica* RCB	Daro_3541
glycosyl transferase A (GTA)	*Dechloromonas aromatica* RCB	Daro_3543
dgkA	*Dechloromonas aromatica* RCB	Daro_3544
cyclic nucleotide-binding domain (CNBD)	*Dechloromonas aromatica* RCB	Daro_3546
ugtP	*Dechloromonas aromatica* RCB	Daro_3547
metallophopshoesterase (MP)	*Dechloromonas aromatica* RCB	Daro_3548
rpoS	*Dechloromonas* sp. JJ	JJ_00006680
nlpD	*Dechloromonas* sp. JJ	JJ_00006670
pcm	*Dechloromonas* sp. JJ	JJ_00006660
surE	*Dechloromonas* sp. JJ	JJ_00006650
hypothetical protein	*Dechloromonas* sp. JJ	JJ_00006640
Pro tRNA	*Dechloromonas* sp. JJ	JJ_00006630

hypothetical protein	*Dechloromonas* sp. JJ	JJ_00006620
ihfA	*Dechloromonas* sp. JJ	JJ_00006610
pheT	*Dechloromonas* sp. JJ	JJ_00006600
pheS	*Dechloromonas* sp. JJ	JJ_00006590
rpoS	*Azospira suillum* PS	Dsui_1971
nlpD	*Azospira suillum* PS	Dsui_1970
pcm	*Azospira suillum* PS	Dsui_1969
surE	*Azospira suillum* PS	Dsui_1968
hypothetical protein	*Azospira suillum* PS	Dsui_1967
Pro tRNA	*Azospira suillum* PS	Dsui_1966
hypothetical protein	*Azospira suillum* PS	Dsui_1965
ihfA	*Azospira suillum* PS	Dsui_1964
pheT	*Azospira suillum* PS	Dsui_1963
pheS	*Azospira suillum* PS	Dsui_1962
gloA	*Dechlorobacter hydrogenophilus* LT-1	LT1_00028180
sygB	*Dechlorobacter hydrogenophilus* LT-2	LT1_00028140
glyQ	*Dechlorobacter hydrogenophilus* LT-3	LT1_00028130
gmhB	*Dechlorobacter hydrogenophilus* LT-4	LT1_00028150
plsC	*Dechlorobacter hydrogenophilus* LT-5	LT1_00028160
DUF45	*Dechlorobacter hydrogenophilus* LT-6	LT1_00028170
hypothetical protein A (*hypA*)	*Dechlorobacter hydrogenophilus* LT-7	LT1_00028570
glycosyl transferase A (GTA)	*Dechlorobacter hydrogenophilus* LT-8	LT1_00028600
dgkA	*Dechlorobacter hydrogenophilus* LT-9	LT1_00028610
cyclic nucleotide-binding domain (CNBD)	*Dechlorobacter hydrogenophilus* LT-10	LT1_00028620
ugtP	*Dechlorobacter hydrogenophilus* LT-11	LT1_00028630
metallophopshoesterase (MP)	*Dechlorobacter hydrogenophilus* LT-12	LT1_00028640
cld	*Ideonella dechloratans*	CAC14884.1
cld	*Alicycliphilus denitrificans BC*	G8QKX3_AZOSU
cld	'Candidatus Nitrospira defluvii'	NIDE1387
narG	*Halogeometricum borinquense* PR3	E4NV35

narG	*Halorubrum lacusprofund*i ACAM34	B9LNI9
narG	*Haloarcula marismortui* VKM B-1809	F2Z6I3
napC	*Magnetospirillum magneticum* AMB-1	Q2W3T5_MAGSA
napC	*Leptothrix cholodnii* SP-6	B1Y6A2_LEPCP
napC	*Bordetella petrii*	A9I7N8_BORPD
nirT	*Magnetospirillum bellicus* VDY	MagVDY_00831
nirT	*Azospirillum* sp. TTI	AzoTTI_07353
nirS1	*Magnetospirillum bellicus* VDY	MagVDY_00830
nirS3	*Magnetospirillum bellicus* VDY	MagVDY_00842
nirS2?	*Magnetospirillum bellicus* VDY	MagVDY_00835-00836
ccmA	*Azospira suillum* PS	Dsui_1167
ccmB	*Azospira suillum* PS	Dsui_1168
ccmC	*Azospira suillum* PS	Dsui_1169
ccmD	*Azospira suillum* PS	Dsui_1170
ccmE	*Azospira suillum* PS	Dsui_1171
ccmF	*Azospira suillum* PS	Dsui_1172
ccmG	*Azospira suillum* PS	Dsui_1173
ccmH	*Azospira suillum* PS	Dsui_1174
ccmI	*Azospira suillum* PS	Dsui_1175
napQ	*Azospira suillum* PS	Dsui_1176
napO	*Azospira suillum* PS	Dsui_1177
napQ	*Azospira restricta* SUA2	AzoSUA2_01063
napQ	*Propionivibrio militaris* MP	PropMP_00017840
yedY	*Magnetospirillum bellicus* VDY	MagVDY_01031
yedY	*Magnetospirillum* sp. WD	MagWD_00008110
yedY	*Magnetospirillum magneticum* AMB-1	YEDY_MAGSA
yedY	*Magnetospirillum* sp. SO-1	H261_05784