Microbial Evolution and the Rise of Oxygen: the Roles of Contingency and Context in Shaping the Biosphere through Time

THESIS BY

Lewis Michael Ward

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy in Geobiology

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2017

(Defended 10 May 2017)

© 2017

Lewis Michael Ward
ORCID: 0000-0002-9290-2567

Preface

What was the Earth like before us? This is a question with many answers: only one of them can be true, but we may never have sufficient information to choose between competing hypotheses. We can only collect more and more information to try to decide between them, narrowing the field, but always with far more plausible anwers remaining than we would like.

When we think about recent history, there are questioned that are answered simply: who won the US presidential election in 1960? Nearly every schoolchild could tell you that it was John F. Kennedy. We have written records, we have videos, we have parents and grandparents who remember election day. The records we have of the event are many, they are consistent with one another, and they are easily interpretable without translation. But as we go farther back in time, these records become more sparse—for the election of 1860, there are written records, but no video or living memory—enough to answer the question, but less robust to nuance and detail. If we are to go farther back—how were the Egyptian pyramids built? We have no confirmed answer at all. We have a range of plausible options, centered around primitive engineering and slave labor, but we also have vociferous proponents of alien technology playing a role. The farther back we go, the more sparse and tenuous our records, the more challenging they are to read, the less certain we are of any given solution and the more room there is for assumption, extrapolation, and postulation.

It is no wonder, therefore, that investigation of the early history of the Earth has tremendous room for error and assumption of answers with insufficient evidence. How simple the architecture of the pyramids seems, when we need to understand the architecture

of biological structures over one million times older! We could answer the question of the construction of the pyramids if we were to uncover a new record of the time—never mind a miraculous video recording of the construction, even some quotidian memorandum sent to the pharaoh detailing the day's efforts to construct his tomb would be revolutionary. In many cases, such records may exist, and the stumbling block is in learning to read them. The key discovery is therefore of a Rosetta Stone, a way to couple an uninterpretable record to one that we already know how to read, which will reveal the meaning of this record, and, potentially, teach us how to read it and allow us to translate other instances written in the same language.

Similarly, if we could uncover new records of life on the early Earth, we could begin to solve longstanding questions of how life and the Earth functioned long, long ago. As luck would have it, there does exist a largely untapped record of life on Earth—the biological record preserved in the genomes and biochemistry of extant organisms. With the application of an appropriate Rosetta Stone we can learn to correlate this new record to information gleaned from the rock record.

As it happens, we are living through the single largest revolution in the study of biology at least since the discovery of DNA nearly a century ago. The technological innovations going into the high-throughput sequencing of DNA—starting with the human genome, and now continuing with genomes of tens of thousands of other species we can grow in culture, and hundreds of thousands more extracted from the environment—have drastically increased the abundance of data available to understanding living systems. These data not only allow us to understand our own genetics and evolution, but can also be used to carefully craft and answer questions about the evolution of life in the recent past

and as far back as the origin of life. This application of this newly available, genomic, biological record of life to the understanding of the early evolution of life on Earth is only possible with carefully curated interdisciplinary understanding of biology—genetics, biochemistry, and especially evolution—coupled to an understanding of geology—the nature and history of the atmosphere, oceans, and the rock cycle, and the geochemical and geological records that preserve signals over billions of years.

It is my goal to utilize these coupled records of rocks and life to develop a more complete understanding of the history and evolution of life, and how its reciprocal interactions and modifications of the Earth and environment over timescales of millions to billions of years.

A crucial aspect of having complementary records of historical events is the ability to confirm the statements and suppositions of one record by way of an independent account in a second or third. If one were to consult certain conspiracy-oriented websites as records of ancient Egypt, one would be left with the firm belief that the pyramids were, in fact, constructed by aliens. Yet if one were to consider alternative records—an encyclopedia, say, or the works of a respected Egyptologist—one would come to a very different conclusion. Our understanding of the early Earth is plagued by similar problems. In a field with so little primary data, there are not only no written records, but even the records we have in the form of sedimentary rocks, geochemistry, and the vestiges of early life in the genomes of organisms found today, have been altered and eroded in the intervening billions of years.

The biological and geological records have each been altered and degraded over time in different ways, but in different ways. Reading the geological record is like reading

an original copy of Herodotus that has been buried in the desert for over 2000 years, exposed to wind, the sun, and the occasional rain. Parts are faded, others are smeared, and other sections are missing completely. Meanwhile reading the biological record is like reading a copy of the Iliad, which has been more-or-less faithfully copied over, hand to hand, mouth to mouth, for just as long—more of it is intact, but small variations have been inherited at every retelling, and so reconstructing the original word-for-word is all but impossible. Or, rather, the biological record is a palimpsest, with the original version somewhere buried underneath millennia of revisions, scribblings, annotations, and other gradual modifications. Both the geological and biological records retain information, preserved from the same period of history, preserving aspects of the same story but telling it in different ways and with distinct biases and both missing crucial sections. And in many cases these records of Earth history are so incomplete that they remain open to interpretation. The same text, read by a dozen individuals, with their own assumptions, own interpretive frameworks, and own attempts at filling in the innumerable gaps, will come to at least a dozen different conclusions about the author, the subject matter, and the history since the original writing. All are hypotheses, none of which are likely entirely correct, but some of which are likely to be less wrong than others. It is our job to choose between all of these interpretations so as to best understand what the early Earth was like and how it has changed since. By combining multiple complementary records, we can fill in the missing pages and solve the outstanding mysteries in the history of life and the Earth.

ACKNOWLEDGEMENTS

The GPS Division at Caltech is a truly special place. I can't imagine a better place to be a scientist. The level of support, intellectual stimulation, and degree of collaboration is unlike anywhere else I have ever been. I've spent the past five years surrounded by fantastic teachers, mentors, colleagues, and friends. As a result, there's almost no one in the department who doesn't deserve some credit for making my time at Caltech so wonderful and so I won't manage to thank everyone individually here. So even if you don't see your name here, thank you. You're awesome.

First and foremost I want to thank my advisor, Woody Fischer. Woody has been so incredibly supportive of me throughout my time at Caltech, and has always encouraged me to pursue crazy research ideas and opportunities for travel and fieldwork. Woody has somehow taught me to think both broadly and deeply about problems in geobiology, spanning ridiculous scales of time and size. I don't know of anyone else who could teach me to be comfortable talking about the size of the biosphere 3 billion years ago one moment and the flow of electrons through a single cell the next. Woody is excited by a broader range of scientific ideas and questions than anyone else I've ever met, and the resulting range of conversations I've had with him about every topic under the sun has certainly made me a better scientist. Working with Woody has been an incredible experience, and I can't imagine a better match of advisor. Even more, working with Woody has been a lot of fun.

I also want to thank Joe Kirschvink, my academic advisor and committee chair. He's been an incredible supporter throughout my time at Caltech (even the parts when we only saw each other in Tokyo). Many of my best Caltech experiences have been thanks to Joe, from fieldwork in Antarctica, to shenanigans in Palau, to every 136 trip. Joe has made sure that I balance my time staring at computer screens with time spent outside and actually staring at rocks, and that has been one of the best things for my geology education as well as my mental health. The opportunities for fieldtrips are one of my very favorite things about GPS, and Joe has played a crucial role in keeping that alive.

Many thanks to Shawn McGlynn, with whom I've learned how awesome hot springs can be. My time working with Shawn in Japan has been an excellent series of adventures, and we even managed to do some good science.

And thanks also to the other members of my committee, Victoria Orphan and Jared Leadbetter, who have helped me keep my research and interests grounded in actual microbiology rather than drifting entirely into unconstrained musing. I also want to thank all the other faculty members I've worked with, taken classes from, and gone on field trips with, including Alex Sessions and John Grotzinger. I've learned so much in my time at Caltech about so many topics that sometimes I can feel it leaking out of my ears, and I'll always be grateful for the time I've spent learning from you.

Jim Hemp was a major influence on my research at Caltech, and has helped shape much of my thinking about the evolution of photosynthesis, respiration, and other aspects microbial

metabolism. I'm lucky to have joined the lab alongside such a great friend and mentor. Thanks also to my other collaborators, including Patrick Shih, Vlada Stamenkovic, and Birger Rasmussen. We've come up with some crazy ideas, and it's been a great time the whole way.

To the members of my first year Pit, Max Lloyd, Daven Quinn, Jen Buz, Mathieu Lapotre, Hayden Miller, and Brooke Dallas—you guys are the best cohort I could possibly hope for. Over the course of the past five years you've been my classmates, roommates, travel buddies, and best friends, and I'm lucky to have shown up here alongside you.

To the other members of the Fischer group past and present, my officemates, and other friends in the department—particularly Sean Mullin, Elle Chimiak, Elizabeth Trembath-Reichert, Jena Johnson, Sarah Slotznick, Kyle Metcalfe, and Usha Lingappa—thank you for everything you've taught me, all of the great conversations, and all of the fun times.

And my experience wouldn't have been possible without the help of administrative and support staff, especially Liz Miura Boyd, Alice Oh, and Julie Lee, who have done so much behind the scenes to make my time at Caltech better and easier.

Thanks to Morgan Cable and the Songam Space Camp crew. Morgan was the first Caltech friend I ever made, and she's been one of my biggest supporters ever since. I'm also deeply appreciative that she brought me on board for Space Camp, which has been an incredibly experience and reminded me about how fun and exciting science and teaching should be.

The International Geobiology Course was a totally foundational experience, and much of this dissertation has grown out of interests and skills I developed on the course. The course also introduced me to some of my best friends in the geobiology community, and as a result every conference I've been to since has been way more fun. I'm deeply appreciative of the time spent on the course, and everyone who contributed to making it happen, especially Frank Corsetti, John Spear, Ann Close, and Amber Brown.

ABSTRACT

We are shaped by our environment, but we then shape it in turn. This interplay between life and the Earth, and how these interactions have shaped both parties through time, is the heart of the discipline of geobiology. My research is fundamentally motivated by a desire to understand how life and the Earth have changed together through time to reach the state that they're at today, and to understand from this history how the coevolution of planet and life may be different on other worlds. The focus of my work has been on how the structure and productivity of the biosphere across time and space has been shaped by the metabolic opportunities provided by the environment—as a result of both biotic and abiotic factors—and the metabolic pathways that are available to life, as a result of evolutionary contingency in the evolution of pathways and their inheritance and horizontal transfer.

The biosphere on Earth today is incredibly productive due to the coupled dominant metabolisms of oxygenic photosynthesis and aerobic respiration, yet these can't always be assumed to have been present—considering life more broadly, for instance in the context of the early Earth and other planets, we have to grapple with how evolutionary contingency and planetary environments interact to constrain the metabolic opportunities and rates of productivity available to the biosphere. In this dissertation, I broadly consider how the size and structure of Earth's biosphere has changed through time as surface environments evolve and metabolic innovations accumulate. These investigations make use of information gleaned from the rock record of the early Earth, as well as the biological record of the history of life as preserved in the genomes, biochemistry, and ecology of extant organisms. These coupled records provide opportunities for constraining estimates of the opportunities for life throughout Earth history and elsewhere in the universe.

PUBLISHED CONTENT AND CONTRIBUTIONS

- Chapter 1:
    - Ward, LM, B Rasmussen, and WW Fischer. Electron donor limitation of the biosphere before oxygenic photosynthesis. In preparation.
        - LMW conceived of the project, designed the study, prepared data and analysis, and wrote the manuscript.
- Chapter 2:
    - Ward, Lewis M., Joseph L. Kirschvink, and Woodward W. Fischer. "Timescales of oxygenation following the evolution of oxygenic photosynthesis." Origins of Life and Evolution of Biospheres 46.1 (2016): 51-65. DOI: 10.1007/s11084-015-9460-3
        - LMW conceived of the project, designed the study, prepared data and analysis, and wrote the manuscript.
- Chapter 3:
    - Ward, LM, P Shih, J Hemp, SE McGlynn, and WW Fischer. A Complex History of Phototrophy Revealed by Novel Chloroflexi Lineages. In preparation.
        - LMW conceived of the project, designed the study, prepared data and analysis, and wrote the manuscript.
- Chapter 4:
    - Ward, LM, A Idei, S Terajima, T Kakegawa, WW Fischer, and SE McGlynn. Microbial diversity and iron oxidation at Okuoku-hachikurou

Onsen, a Japanese hot spring analog of Precambrian iron formation. *Geobiology*, in revision.

- LMW conceived of the project, designed the study, prepared data and analysis, and wrote the manuscript.

- Chapter 5:
  - Ward, LM, Stamenković V, Fischer WW. Follow the Oxygen: Comparative Histories of Planetary Oxygenation and Opportunities for Life. In preparation.
    - LMW conceived of the project, designed the study, prepared data and analysis, and wrote the manuscript.
- Appendix 1:
  - Ward, LM, J Hemp, and Woodward W. Fischer. A reduced, abiotic nitrogen cycle before the Rise of Oxygen. In prep.
    - LMW conceived of the project, designed the study, prepared data and analysis, and wrote the manuscript.
- Appendix 2:
  - **Ward, LM,** WW Fischer, K Matsuura, and SE McGlynn. Cone-forming microbial mats of Nakabusa Onsen, Japan as analogs of Precambrian stromatolites. In prep.
    - LMW conceived of the project, designed the study, prepared data and analysis, and wrote the manuscript.

- Appendix 3:
    - **Ward, LM,** A Idei, WW Fischer, and SE McGlynn. Microbial diversity and productivity of Jinata Onsen, an iron-rich intertidal hot spring in Japan. In prep.
        - LMW conceived of the project, designed the study, prepared data and analysis, and wrote the manuscript.
- Appendix 4:
    - **Ward LM**, Hemp J, Pace LA, Fischer WW. 2015. Draft genome sequence of *Leptolinea tardivitalis* YMTK-2, a mesophilic anaerobe from the *Chloroflexi* class *Anaerolineae*. Genome Announc 3(6):e01356-15. DOI: 10.1128/genomeA.01356-15
        - LMW analyzed the data and wrote the manuscript.
    - **Ward LM**, Hemp J, Pace LA, Fischer WW. 2015. Draft genome sequence of *Herpetosiphon geysericola* GC-42, a nonphototrophic member of the *Chloroflexi* class *Chloroflexia*. Genome Announc 3(6):e01352-15. DOI: 10.1128/genomeA.01352-15
        - LMW analyzed the data and wrote the manuscript.
    - Hemp J, **Ward LM**, Pace LA, Fischer WW. 2015. Draft genome sequence of *Levilinea saccharolytica* KIBI-1, a member of the *Chloroflexi* class *Anaerolineae*. Genome Announc 3(6):e01357-15. DOI: 10.1128/genomeA.01357-15
        - LMW analyzed the data and wrote the manuscript.

- Pace LA, Hemp J, **Ward LM**, Fischer WW. 2015. Draft genome of *Thermanaerothrix daxensis* GNS-1, a thermophilic facultative anaerobe from the *Chloroflexi* class *Anaerolineae*. Genome Announc 3(6):e01354-15. DOI: 10.1128/genomeA.01354-15
    - LMW assisted in analysis of the data and writing of the manuscript.
- Hemp J, **Ward LM**, Pace LA, Fischer WW. 2015. Draft genome sequence of *Ornatilinea apprima* P3M-1, an anaerobic member of the *Chloroflexi* class *Anaerolineae*. Genome Announc 3(6):e01353-15. DOI: 10.1128/genomeA.01353-15
    - LMW analyzed the data and wrote the manuscript.
- Hemp J, **Ward LM**, Pace LA, Fischer WW. 2015. Draft genome sequence of *Ardenticatena maritima* 110S, a thermophilic nitrate- and iron-reducing member of the *Chloroflexi* class *Ardenticatenia*. Genome Announc 3(6):e01347-15. DOI: 10.1128/genomeA.01347-15
    - LMW assisted in analysis of the data and writing of the manuscript.

- Appendix 5:
    - Shih, P., Hemp, J., **Ward, L.**, Matzke, N. & Fischer, W. 2016. Crown group oxyphotobacteria postdate the rise of oxygen. *Geobiology* 15.1 (2017): 19-29. DOI: 10.1111/gbi.12200
        - LMW assisted in conceiving of the study, analysis of the data, and writing of the manuscript.

- Appendix 6:
    - Stamenković, V, LM Ward, M Mishna, and WW Fischer. Aerobic Environments on Mars. *Nature*, submitted.
        - LMW assisted in conceiving of the study and writing of the manuscript.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Introduction

One of Darwin's great dilemmas when he published the Origin of Species was the sudden appearance of diverse animals at the beginning of the known fossil record, seemingly out of nowhere—what we now call the Cambrian Explosion, starting around 542 million years ago. He hypothesized that further investigation into the early rock record would reveal earlier, simpler forms that gave rise to the appearance of animals. However, it has only been in the past few decades that Darwin's suspicions have been proven absolutely correct. Not only was there an earlier history of simpler animals leading up to the Cambrian Explosion, but there is also a record of microbial life going back more than 3.7 billion years ago, almost to the beginning of the rock record on Earth.

The modern diversity of complex organisms like plants and animals is therefore a relatively recent innovation. The history of microbes like bacteria and archaea is much longer, spanning a period of time so long that the entire evolutionary history of animals—from simple sponge-like organisms to the entire diversity of jellyfish, bees, giant squid, and humans—could have occurred half a dozen times with hundreds of millions of years to spare.

It is therefore no wonder that the history of microbial evolution is a rich, complicated story, albeit a difficult one to read. While the history of animal life is recorded in fossils with characteristic morphologies, the bacterial fossil record is sparse and largely uninformative. Bacteria are not easily preserved, and those microfossils that do exist are nearly impossible to confidently classify because bacterial morphology is simple and similar forms can evolve convergently in distantly related taxa. If we are to understand the

evolutionary history of bacteria and other microbes in deep time, we need to identify new records of past life and develop new strategies to investigate them.

The most important stories in microbial evolution are not based on morphology anyway, but on metabolism. One of the most significant innovations in the history of life was the invention of oxygenic photosynthesis by Cyanobacteria. Oxygenic photosynthesis, the biological ability to oxidize water using light in order to fix organic carbon, produces O_2 as a brproduct, and led to the oxygenation of Earth's atmosphere ~2.3 billion years ago. Oxygenic photosynthesis made complex, multicellular life possible, and revolutionized geochemical cycles and the metabolic opportunities available for life by providing molecular oxygen for oxidative weathering and aerobic respiration.

While animals all make a living by breathing oxygen and eating organic carbon, microbes can possess any number of a bewildering variety of metabolisms—from eating hydrogen to produce methane, to breathing heavy metals like uranium, to conserving energy by using light to rust iron. Many of these microbial metabolisms only exist in niche environments today, but are vestiges of earlier periods in Earth history—particularly before the oxygenation of the atmosphere ~2.3 billion years ago—when they were major parts of the biosphere. These metabolisms, and the organisms that possess them, present us with valuable targets for investigating the evolution of metabolism and ways to understand the ancient biosphere.

The nature of the biosphere before the rise of oxygen—the metabolisms that existed, the amount of productivity they supported, and the structure of biogeochemical cycles—is one of the largest and most exciting open questions in historical geobiology. However,

there is much uncertainty in the metabolisms that were possible, what metabolisms had evolved by this time, how efficient they are at driving productivity, and how this played in to other Earth system characteristics such as the composition of the atmosphere. If these questions can be answered, we can better understand the history of our own planet and how it came to be the place we live today, and we can also make predictions about what biospheres on other planets may be like and how to detect them if they do exist. The first section of this dissertation therefore addresses the nature of the early Earth, before the evolution of oxygenic photosynthesis and the Rise of Oxygen. These chapters primarily utilize geochemical models of the flow of electrons between redox-active molecules utilized by biology, including electron donors like molecular hydrogen and ferrous iron supplied by geology but used by early life to support primary productivity. These box models are relatively simple, utilizing previous estimates of the fluxes of geological and atmospheric processes on the early Earth to make order of magnitude estimates of the scale of biological productivity. Because these models are relatively simple and grounded in the best available constraints from the rock record, they are relatively insensitive to uncertainties in poorly constrained aspects of the early Earth like volcanic outgassing rates or the size and distribution of continents. This is in contrast to other, much more complex models, which can have unexpected nonlinear dependencies on poorly constrained parameters. These models provide valuable predictions about the relative scales and rates of biological processes as related to geological ones, and introduce new ways of thinking about the relationship between early life and the Earth.

Chapter 1 considers the potential rates of primary productivity that could be supported by the early biosphere, dominated by methanogens and/or anoxygenic phototrophs, and concludes with estimates of productivity many orders of magnitude lower than are characteristic of the modern, oxygenic photosynthesis-fueled biosphere. Chapter 2 considers the timescales over which life and the Earth change in response to the evolution of oxygenic photosynthesis, with the result that oxygen accumulates rapidly and therefore the GOE should follow closely the evolution of oxygenic Cyanobacteria. Ultimately, this section paints a picture of the pre-oxygen biosphere as being much smaller and less productive than what we see today, and highlights the significance of oxygenic photosynthesis—and the ability to use water as a virtually unlimited electron donor—as the fundamental driver of the large, complex, productive biosphere we see today that allows life to develop geologically-relevant feedbacks for maintaining climate, atmospheric composition, and biogeochemical cycles.

Insights from the rock record, and geochemical models like those described in the first three chapters of this dissertation, are valuable, but are limited by the incompleteness of the rock record due to billions of years of gradual degradation due to weathering and tectonics. To understand how the early biosphere functioned, we need to develop novel techniques for supplementing the incomplete rock record with the other record of life on Earth: the biological record, as preserved in the diversity of extant organisms and their genomes. By developing and testing hypotheses using both the rock and biological records, we can develop a much more thorough understanding of how life functioned on the early Earth. In Chapter 3, I investigate the biological record of anoxygenic phototrophy in the Chloroflexi

phylum of bacteria to address the antiquity of anoxygenic photosynthesis. I demonstrate from comparative phylogenetics that the distribution of phototrophy in this clade is the result of horizontal gene transfer, and therefore may be a relatively late evolutionary occurrence. This is an important result for understanding not only the history of this clade, but also the role of horizontal gene transfer in shaping the diversity and distribution of metabolic pathways across the tree of life. Furthermore, by demonstrating that anoxygenic photosynthesis in the Chloroflexi may be a relatively late innovation, it helps to demonstrate that in many cases it may not have been crown group playing a role in driving the early, pre-oxygen biosphere. Instead, the Archean biosphere was likely dominated by extinct stem group lineages, and that these may have been wiped out during adaptive radiations following the expansion of productivity and new metabolic and ecological opportunities following the GOE.

An alternative strategy for understanding the ancient Earth is by characterizing niche environments that support processes that are in some ways analogous to those that occurred on the ancient Earth. These process analog environments can reveal the interplay of environmental conditions and microbial communities that determine rates of primary productivity, the structure of geochemical cycles, and preservable geological and geochemical signatures. In Chapter 4 of this dissertation, I characterize a novel iron-rich hot spring microbial community in Japan that is in some ways analogous to periods in Earth history. This hot spring, Oku-okuhachikurou Onsen in Akita Prefecture, Japan, supports aerobic iron oxidizing bacteria and deposits highly oxidized, organic carbon-lean iron deposits similar to Banded Iron Formations deposited after the Rise of Oxygen. This

system highlights the low metabolic yield per electron donor of lithotrophic bacteria relative to phototrophs, a largely underappreciated but potentially far-reaching trend that may have relevance for ancient Earth environments described in Chapter 1 and diverse planetary environments discussed in Chapter 5.

Stephen J. Gould once remarked that if one were to rewind the tape of life and play it back from some earlier point, there is no telling if it would play out the same or if the diversity of life would turn out wildly different. The biosphere we have today is the result of billions of years of stochastic phenomenon, contingent on the internal randomness of evolution as well as external influences from the environment. On other worlds, where the environment may provide very different metabolic opportunities for life, the trajectory of evolution may play out very differently. Understanding how Earth's biosphere evolved to reach its present state provides a framework for predicting how life may develop on other worlds with very different conditions. I conclude this dissertation with a consideration of how the history of oxygen on Earth differs from that on Mars, Europa, and other planets, and what these comparative histories may mean for the opportunities for life on other worlds. This chapter highlights the role that planetary context plays in shaping the metabolic opportunities available to life, and how this may relate to the evolutionary trajectories that life can follow. On the early Earth, life was limited to making use of limited geological fluxes of electron donors to drive low energy and low productivity metabolisms until the evolution of oxygenic photosynthesis freed the biosphere from these constraints. On other planets, the balance of electron donors and acceptors available to life

may vary, resulting in some environments where less metabolic evolution is necessary to achieve highly energetic metabolisms than on Earth.

Finally, this dissertation includes several appendices, including projects on which I was a substantial contributor but not lead author and projects that are underway but not yet complete. Appendix 1 considers the nitrogen cycle on the early Earth, before the rise of oxygen, and demonstrates that in the absence of molecular oxygen it was likely simple, restricted to reduced compounds, and primarily driven by abiotic processes. This ties in to the first section of the dissertation in demonstrating that before the evolution of oxygenic photosynthesis, the biosphere may not have been capable of driving geochemical cycles as it does today. Appendix 2 discusses a hot spring ecosystem supporting cone-forming microbial mats analogous to those interpreted that produced conical stromatolites in the ancient rock record. Stromatolites are frequently interpreted as evidence of oxygenic Cyanobacteria. This analog system however supports cone-shaped mat morphologies developed by anoxygenic Chloroflexi, providing good evidence that diverse microbial taxa can lead to similar mat and stromatolite morphologies, and therefore ancient stromatolites may have formed in the absence of oxygenic Cyanobacteria. In Appendix 3, I describe Jinata Onsen, on Shikinejima Island, Japan, where an iron- and hydrogen-rich hot spring mixes with sulfate- and oxygen-rich seawater along its outflow, and so over short spatial scales recapitulates the range of conditions present in Earth's oceans over the Proterozoic eon, in which the oceans transitioned from oxygen-poor and iron-rich to modern conditions low in iron but rich in sulfate and oxygen. Appendices 2 and 3 tie in closely to Chapters 3 and 4; these appendices describe the field sites from which the metagenomes analyzed in

Chapter 3 were sampled, and the use of these field sites as analogs of ancient environments is complementary to that described in Chapter 4. Appendix 4 describes genomic sequencing efforts of diverse nonphototrophic Chloroflexi, expanding representation of sequenced members of this phylum to span all of the described classes, and uncovering previously unrecognized metabolic diversity. This ties in closely to Chapter 3, and together these two sections greatly improve our understanding of the metabolic diversity of previously understudied lineages of Chloroflexi. Appendix 5 discusses the evolution of oxygenic photosynthesis in Cyanobacteria, using cross-calibrated relaxed Bayesian molecular clocks to date the radiation of crown group oxygenic Cyanobacteria ("Oxyphotobacteria") and their divergence from their recently-discovered nonphototrophic sister clade Melainabacteria. This analysis demonstrates that oxygenic photosynthesis must be fairly young, less than ~2.5 billion years old, consistent with estimates for the timing of the GOE estimated in Chapter 1. This study also demonstrates that the crown group radiation of Oxphotobacteria occurred ~2 Gya, and that the GOE must therefore have been driven by stem group lineages, highlighting the importance of extinct and stem lineages in powering the ancient Earth. Appendix 6 is a companion to Chapter 5, and describes innovative modeling of oxygen solubility in low-temperature liquid brines on Mars, demonstrating that under a range of Martian conditions today and in the past oxygen may accumulate to substantial concentrations, high enough to support respiration by bacteria and potentially even simple multicellular animals like sponges.

Chapter 1

ELECTRON DONOR LIMITATION OF THE BIOSPHERE BEFORE THE ORIGIN OF OXYGENIC PHOTOSYNTHESIS

Ward, Lewis M., Birger Rasmussen, and Woodward W. Fischer. In preparation.

Abstract:

The evolution of oxygenic photosynthesis, and the resulting oxygenation of the atmosphere ~2.3 billion years ago, was the single largest shift in the size and structure of the biosphere since the origin of life. Oxygen-fueled metabolisms are highly energetic, and drove the evolution of complex life in Neoproterozoic and Phanerozoic time, and oxygenic photosynthesis itself came to dominate primary productivity on the surface of the Earth. Perhaps more revolutionary than the introduction of molecular oxygen, oxygenic photosynthesis introduced the ability for life to utilize water—an essentially infinite electron donor—for autotrophy. Before the evolution of oxygenic photosynthesis, the biosphere was supported by anaerobic metabolisms such as methanogenesis, acetogenesis, and anoxygenic photosynthesis. In addition to being less energetic than oxygen-based metabolisms, these early autotrophic processes depend on geologically supplied electron donors such as molecular hydrogen and ferrous iron as a source of electrons for reducing CO_2 to organic carbon. As a result, the geological fluxes of electron donors may have served as a fundamental limitation to the productivity of the early biosphere, rather than nutrients like phosphate or fixed nitrogen as today. Here, we integrate estimates for the geological fluxes of electron donor compounds, phosphate, and fixed nitrogen to determine the fundamental limits to pre-oxygen productivity. We demonstrate that even if iron

oxidation was occurring, it was a relatively minor fuel for productivity compared to molecular hydrogen. Furthermore, we show that fluxes of fixed nitrogen and phosphate were substantially in excess of rates of productivity allowed by the supply of electron donor compounds, even in the absence of biological nitrogen fixation. This suggests that until the evolution of oxygenic photosynthesis, the size and productivity of the biosphere were fundamentally constrained by the geological supply of electron donors, and that there may not have been evolutionary pressure to evolve biological nitrogen fixation until after the evolution of oxygenic photosynthesis.

Introduction:

While the antiquity of oxygenic photosynthesis by Cyanobacteria is a matter of significant debate (Johnson et al. 2013a, Johnson et al. 2013b, Crowe et al. 2013, Planavsky et al. 2014, Lyons et al. 2014, Fischer et al. 2016, Shih et al. 2016, Ward et al. 2016), molecular oxygen is constrained to extremely low concentrations earlier in Earth history (Farquhar et al. 2000, Johnson et al. 2013), but rapidly accumulated to significant concentrations at the GOE (Goldblatt et al. 2006, Bekker and Holland 2012, Ward et al. 2016, Luo et al. 2016).

Perhaps the most significant impact of oxygenic photosynthesis was not the introduction of O_2 into biogeochemical cycles, but rather the ability for life to make use of an essentially unlimited electron donor—water. The fixation of inorganic carbon into biomass fundamentally involves the reduction of CO_2 into more reduced forms via the transfer of electrons from an electron donor. Water as an electron donor for oxygenic photosynthesis is available in an essentially unlimited supply, and so rates of primary

productivity today are constrained by other factors such as the availability of nutrients iron, fixed nitrogen, and phosphorous (Tyrrel 1999). Before the evolution of oxygenic photosynthesis—whenever this metabolism evolved, whether it was near the GOE or much deeper in Archean time—biology was restricted to more limited electron donor compounds, potentially including molecular hydrogen, ferrous iron, and reduced sulfur compounds (e.g. Kharecha et al. 2005, Canfield et al. 2006). These electron donors are geologically sourced and typically available at much lower abundances than water, raising the possibility that electron donors—rather than other nutrients like phosphate and fixed nitrogen—limited rates of primary productivity on the early Earth, before the evolution of oxygenic photosynthesis.

Here, we consider the contributions of electron donor compounds—particularly molecular hydrogen and ferrous iron—to fueling productivity on the pre-oxygen Earth, and the supply of these electron donors relative to nutrients like nitrogen and phosphorous. We conclude that nutrients, including phosphate and fixed nitrogen, were likely available in excess of the geological supply of electron donor compounds. This suggests that, to first order, the supply of electron donor compounds like H_2 set the tempo of early productivity, and that the resulting rate of productivity was at least 1000x lower than today. Furthermore, the efficiency with which these electron donors were consumed would be a major constraint to productivity, as consumption of H_2 via different metabolisms can affect the yield of organic carbon by an order of magnitude. We also conclude that at least at a global scale, abiotic nitrogen fixation pathways may have been sufficient to meet the nitrogen demands of the biosphere, and so there may not have been

evolutionary pressure to evolve biological nitrogen fixation until the evolution of oxygenic photosynthesis. The relative abundance of nutrient fluxes relative to electron donors also suggests that there may have been an accumulation of nutrients in the pre-oxygen ocean, potentially providing a massive boost to productivity immediately following the evolution of oxygenic photosynthesis, leading to the rapid accumulation of O_2 at the GOE. The GOE therefore represents a singular break in the uniformitarian history of the scale of life on Earth, in which the tempo of productivity shifted geologically instantaneously form being set by slow geological rates of electron donor supply to being constrained by biological uptake and recycling of nutrients.

What limited productivity of the early biosphere?

In the modern ocean, it is the supply of nutrients such as phosphate and fixed nitrogen that limit primary productivity (Tyrrell 1999). However, this may not have been true on the early Earth, before the evolution of oxygenic photosynthesis. To determine the ultimate limiting factor on pre-oxygen productivity, we must consider the fluxes of nutrients to the Archean ocean, and the rates of productivity they can support (assuming an elemental composition of biomass approximating the Redfield Ratio of 106 C: 16 N: 1 P), compared against the number of electrons available for the fixation of organic carbon from a given flux of electron donor compounds (assuming an oxidation state of organic carbon similar to the modern, around -1, therefore requiring 5 electrons to fix one molecule of CO_2 to one molecule of organic carbon). We have assembled the best available estimates for the fluxes of essential nutrients to the pre-oxygen biosphere, focusing on phosphate, fixed nitrogen, and the electron donors molecular hydrogen and ferrous iron (Table 1). Though

estimates of the availability and fluxes of these compounds are poorly constrained, in many cases relying on unknown conditions such as the composition of the early atmosphere, they are generally acceptable for an order of magnitude estimate—a level of accuracy acceptable for the results that follow.

Major electron donors

Previous estimates of Archean productivity have relied on steady state estimates of dissolved iron content of Archean seawater and atmospheric hydrogen concentrations with models of upwelling or diffusive flux into the photic zone (e.g. Kharecha et al. 2005, Canfield et al. 2006). These models assume steady state concentrations balanced by biological oxidation, recycling, and geological supply. Here, we are interested in the number of electrons delivered from the Earth to life, and the resulting amount of biomass (here considered as moles of organic carbon) that can be produced per year. We therefore use the flux of electron donors like ferrous iron and molecular hydrogen to the oceans/atmosphere, as estimated by geological processes like hydrothermal alteration of crust and volcanic outgassing, without assumption of particular steady state concentrations.

Production of oceanic crust and associated reduced iron was derived from Lécuyer and Ricard (1999). Based on Williams et al. (2012), it is assumed that mantle redox—and therefore the Fe^{2+}/Fe^{3+} ratio of new crustal material—has been constant since accretion. Hydrothermal fluxes were derived from Elderfield and Schultz (1996) for modern values as a reference. It has been argued that Archean hydrothermal fluxes may have been up to ten times higher than modern based on Des Marais (2000), although it is possible that Archean

hydrothermal fluxes were lower than today (e.g., Korenaga 2006); as a result of the uncertainty of even the sign in this difference we use the modern hydrothermal flux.

Molecular hydrogen is supplied to the fluid Earth by a variety of sources, including volcanism and water-rock interactions in the oceans and continental subsurface. Fluxes of volcanic gases including H_2 have been estimated by Holland (2002) and the contribution of various pathways of continental and oceanic water-rock interactions to H_2 fluxes have been collected by Sherwood Lollar et al (2014) and summarized in Table 1.

The fluxes of ferrous iron and molecular hydrogen are estimated here to be of the same order (~1.5-6x10^{12} moles/year). If fluxes of both of these compounds are stoichiometrically oxidized to drive organic carbon fixation via anoxygenic photosynthesis, the contribution of hydrogen to productivity will be double that of iron as the oxidation of H_2 to H^+ provides two electrons to the one form Fe^{2+} to Fe^{3+}. Therefore even in the most productive scenario presented here (Scenario 1, Table 2), hydrogen will fuel several times more productivity than hydrogen.

Phosphate

Phosphate is necessary for all known life, and appears to be the primary limiting nutrient in the modern ocean over long timescales (Tyrrel 1999), and so it availability is crucial for constraining potential rates of primary productivity. However, it is challenging to constrain the availability of phosphate to the early biosphere, as the net abundance of marine phosphate is the result of the interplay of a wide range of geological, hydrological, biological, and other factors. We therefore consider a variety of influences on phosphate

supply, recycling, and burial in order to make conservative estimates of the net availability of phosphate to the pre-oxygen biosphere.

The ultimate source of phosphate to the biosphere is chemical weathering of phosphate minerals in rocks. The extent of chemical weathering occurring on the Archean Earth—and therefore the delivery of phosphate to the oceans—is evident through the extensive production of carbonate minerals, which require a substantial flux of alkalinity (another product of chemical weathering) to the oceans (Higgins et al. 2009).

While chemical weathering of continental crust is the primary mechanism for oceanic phosphate delivery today, weathering and alteration of seafloor basalts—such as at hydrothermal vents—may have been significant on the early Earth. Basalt typically contains much higher phosphorous concentrations than average continental crust (up to ~8000ppm versus 700ppm, Taylor et al. 1995, Horton 2015), and so has the potential to be a major phosphate source to the ocean. Hydrothermal vents in the modern ocean are thought to be a net sink of phosphate (e.g. Elderfield and Schultz 1996), largely as a result of the sorption of phosphate to metal oxides that precipitate as metal-rich hydrothermal fluids interact with oxidized seawater (Wheat et al. 1996). Before the oxygenation of the oceans led to extensive water column metal oxidation, hydrothermal systems may have instead served as a major source of phosphate to the ocean.

Since even the sign of difference in phosphate weathering between the Archean and modern Earth is unclear, we include as a first-order estimate of weathering-based phosphate delivery to the oceans from modern riverine input, around 7×10^{10} moles P/year (Tyrrel 1999) (Table 1). This estimate is approximate over a range of a fewfold, due to

uncertainties in the lithology of the early crust and weathering rates, but is a reasonable starting estimate, particularly given that in the modern ocean most productivity is supported not by new delivery of phosphate, but by recycling. Biology has developed multiple strategies for maximizing uptake and recycling of this vital nutrient, and as a result the relationship between phosphate input and primary productivity in the modern ocean suggests that phosphorous is recycled on the order of 500 times between input to the ocean and burial (Tyrrel 1999). Whether this extent of efficient recycling occurred on the early Earth is an open question, but recycling almost certainly did occur. We therefore provide an additional estimate for phosphate delivery to the photic zone via upwelling of more nutrient-rich bottom waters in a manner analogous to the modern ocean, using estimates of deep-water phosphate concentrations (from phosphate concentrations in marine sedimentary rocks like banded iron formations) and estimates of rates of upwelling and diffusivity. Assuming deep-water phosphate concentrations constrained to about 10% of modern due to sorbtion of phosphate onto iron oxides (Bjerrum and Canfield 2002), along with a typical diffusivity of ~100 m/year, delivery of phosphate to the surface ocean would be on the order of 1.3×10^{12} moles P/year (Table 1). This is a conservative estimate, as the Archean iron cycle may have been dominated by ferrous silicates, not iron oxides (Box 1), and ferrous silicates are not expected to bind phosphate as efficiently as iron oxides.

Fixed nitrogen

In the modern ocean, fixed nitrogen is supplied primarily by biological nitrogen fixation using the nitrogenase enzyme (Canfield et al. 2010). The antiquity of nitrogenase is contested, with estimates ranging from its presence in the last universal common ancestor

to a much later evolution, near the GOE (Weiss et al. 2016, Boyd et al. 2011). Before the evolution of nitrogenase, and potentially throughout the Archean, fixed nitrogen may have been provided primarily by abiotic nitrogen fixation processes (Navarro-Gonzalez et al. 2001). Abiotic nitrogen fixation occurs through a variety of processes, and the fluxes and products of these reactions dependent on atmospheric chemistry and other factors. These include the fixation of N_2 into reduced forms via high temperature and reducing conditions at hydrothermal vents (~10^{10} moles N per year, Brandes et al. 1998), fixation of N_2 to NO (which would reach the oceans as NO_3^-, Mancinelli and McKay 1988) by lightning (~10^{12} moles N per year, Navarro-Gonzalez et al. 1998) and the photochemical production of HCN (~$2x10^{12}$ moles N per year, Tian et al. 2011).

The primary mechanism of abiotic N fixation, and the N species produced, is expected to vary with the CH_4:CO_2 ratio of the early atmosphere. Under a high CO_2 atmosphere, lightning-based N fixation would produce large fluxes of NO, which would reach the oceans as NO_3^- on the order of 10^{12} moles N/year (Mancinelli and McKary 1988, Navarro-Gonzalez et al 1998, Navarro Gonzalez et al 2001). However, under the relatively high methane atmospheric conditions that are expected to have been present throughout the Archean (e.g. Kasting et al. 2001, Pavlov et al. 2001), lightning-based N fixation would have primarily produced HCN (Navarro Gonzalez et al. 2001), and photochemical HCN production may have been quite high (Tian et al. 2012), potentially totaling ~$3x10^{12}$ moles N/yr. Under either scenario, fluxes of fixed nitrogen would have been on the order of 10^{12} moles N per year, at least an order of magnitude in excess of the nitrogen demands of the electron donor-limited biosphere, even without recycling of organic nitrogen, which may

be significant—particularly before O_2 was available to fuel nitrification and a complete nitrogen cycle (Zerkle and Mikhail 2017).

Other nutrients and electron donors

While sulfur isotope analysis of 3.45 Ga stromatolites indicates that microbial sulfate reduction and sulfur disproportionation were active at this time (Bontognali et al. 2012), it is expected that sulfur played only a minor role in the Archean ocean due to the rapid titration of sulfide to pyrite in an iron-rich water column (Canfield 2004). As a result, sulfate concentrations were very low in Archean seawater (Habicht et al. 2002, Crowe et al. 2014), and so the contribution of sulfate reduction to remineralization was likely minor, and recycling of sulfur was minimal. In sum it is therefore likely that sulfur metabolisms contributed only very small amounts of productivity to the Archean biosphere, orders of magnitude less than molecular hydrogen (Kharecha et al. 2005, Canfield et al. 2006). We therefore do not include a separate calculation for the expected productivity of sulfur-based metabolism; incorporation of a sulfur-based productivity model would shift our expected rate of productivity slightly, but because the availability of sulfur is much less than iron (Walker and Brimblecombe 1985), this shift would be even smaller than the difference between an anoxygenic phototrophic biosphere with and without iron oxidation (Scenario 1 and Scenario 2. Assuming a volcanic SO_2 outgassing flux of $\sim 10^{11}$ moles/year (Ono et al. 2003), sulfur could fuel at most only about 4×10^{10} moles carbon fixed/year, less than 2% of our estimate for H_2-fueled productivity.

While it has been proposed previously that the supply of trace metal cofactors may have been a limitation to the productivity of early Cyanobacteria (e.g. Anbar and Knoll

2002, Saito et al. 2003), this has not been demonstrated for the anoxygenic organisms expected to drive primary productivity before the evolution of oxygenic photosynthesis. Consideration of individual trace metals shows that for metals important to methanogens and anoxygenic phototrophs (e.g. Fe, Ni) were relatively abundant early in Earth history relative to metals that are less essential for these organisms (e.g. Cu, Mo, Zn) (Williams and Rickaby 2012, Robbins et al. 2016). These organisms may even have inherited their suite of metal cofactors as a result of their evolution and diversification in the early ferruginous ocean as has been proposed for Cyanobacteria in a euxinic Paleoproterozoic ocean (Saito et al. 2003).

While a source of inorganic carbon is necessary to support autotrophic carbon fixation, this is not expected to have been limiting to the early biosphere. The composition of the Archean atmosphere is poorly constrained, but it is generally thought that pCO_2 was significantly higher than today (e.g. Blättler et al. 2016), partially due to geological constraints such as the mineralogy of Archean paleosols (Rye and Holland 1995) and partially as a modeled solution for maintaining a clement climate under the faint young sun (Kasting 1987). CO_2, as dissolved inorganic carbon, should therefore have been well in excess of the demands of the early biosphere, both as a source of carbon for fixation of biomass as well as the terminal electron acceptor for cellular redox balance in metabolisms like methanogenesis and anoxygenic photosynthesis.

The electron donor-limited early biosphere

Our results suggest that electron donors, of which H_2 was likely most significant, were limiting to productivity before the evolution of oxygenic photosynthesis (Figure 1,

Table 2). Although iron oxidation has been considered a significant source of energy to the early biosphere, its relative contribution to rates of primary productivity is contested (e.g. Kharecha et al. 2005, Canfield et al. 2006), and may have been negligible (Box 1). Molecular hydrogen is thought to have been another major electron donor at this time, capable of fueling not only anoxygenic photosynthesis but also metabolisms like methanogenesis and acetogenesis. Based on our estimates of H_2 and Fe^{2+} fluxes to the pre-oxygen ocean, it appears that molecular hydrogen fluxes are somewhat higher than those of iron (Table 1), making hydrogen a more significant fuel for primary productivity than iron, even if photoferrotrophy oxidizes all of the ferrous iron sourced to the ocean. Global rates of primary productivity were likely no higher than 2.75×10^{12} moles C/yr, perhaps somewhat less if iron were not serving as an electron donor for photosynthesis. This value is more than 1000x lower than modern productivity, and much lower than previous estimates of Archean productivity (e.g. Kharecha et al. 2005, Canfield et al. 2006) (Figure 2). While local environments may have remained limited in other nutrients, on an averaged global scale it is the geological flux of electron donors that fundamentally limited the biosphere. This makes sense, as organisms can evolve to become more efficient with their use and recycling of nutrients, but electron donors are stoichiometrically consumed during the fixation of organic carbon.

Molecular hydrogen is an incredibly versatile and favorable electron donor for diverse metabolisms (hydrogenotrophy), including methanogenesis, acetogenesis, anoxygenic photosynthesis, and both aerobic and anaerobic lithotrophy. Of these, lithotrophy was likely not significant before the rise of oxygen, as it relies on respiratory electron acceptors

such as oxygen or sulfate that were likely scarce or absent. Methanogenesis and acetogenesis are thought to be among the earliest metabolisms to have evolved (Battistuzzi et al. 2004), and so may have been responsible for driving productivity in the earliest stages of life on Earth. Sometime later, anoxygenic photosynthesis evolved. The evolution of anoxygenic photosynthesis may have led to a significant boost in primary productivity, as this metabolism can be more efficient at carbon fixation for a given flux of electron donor. Methanogenic productivity is limited not only by the flux of electron donor, but also limited by the necessity of consuming electron donor to supply cellular energy as well as to fix carbon—a limitation absent in anoxygenic phototrophs that can run cyclic electron flow to generate ATP (Madigan et al. 2012). Methanogens, as a result, must channel electrons into methane to conserve energy, and only fix on the order of 1 mole of organic carbon for every 10 moles of methane generated, dropping their expected productivity by an order of magnitude (e.g. Thauer et al. 2008). This leads to a drop in yield from ~2.4×10^{14} cells/mol H_2 for anoxygenic phototrophs to ~1.7×10^{13} cells/mol H_2 for methanogens (assuming average oxidation state of organic carbon of -1, ~10fg carbon per cell, and maximum yields for methanogens without cytochromes of 3g per mole of CH_4, Whitman et al. 1998, Thauer et al. 2008). As a result of this relative inefficiency of methanogenesis, primary productivity could have increased by about a factor of ten following the evolution and expansion of anoxygenic photosynthesis (from Scenario 2 to Scenario 3 in Table 2 and Figure 1). Determining the age of anoxygenic photosynthesis will therefore be important for plotting primary productivity through time.

Even under a hydrogen-fueled biosphere, iron is the ultimate source of most electrons from the solid Earth—water/rock interactions in hydrothermal vents, continental crust, and other environments result in the transfer of electrons from iron in rocks to molecular hydrogen, which is then released to the biosphere (Sherwood Lollar et al. 2014). Iron is the single largest reservoir of electrons in the Earth's crust (Walker and Brimblecombe 1985), yet is relatively inaccessible to biology, and so the extent to which life can be rock-powered is dependent on the extent of hydrothermal alteration and delivery to the oceans. Water is therefore a crucial intermediary in the transfer of electrons from rocks to life, particularly in the absence of photosynthesis.

Is a low-productivity Archean biosphere consistent with the carbon record?

An important question about the limited biological productivity predicted here is whether this rate of production is consistent with the abundance of organic carbon buried throughout the Archean. It has been suggested that organic carbon burial during the Archean was of similar order to today based on evidence from the organic carbon content of Archean sedimentary rocks and trends in the carbon isotope record (e.g. Lyons et al. 2014). These records, however, are consistent with the rates predicted here when differences in burial rates, and carbon isotope fractionation are taken into account.

In the modern, well oxygenated ocean, organic carbon burial rates are on the order of 0.2% of NPP (Tyrrell 1999); while carbon burial rates on the early Earth are not well constrained (Ward et al. 2016), they were almost certainly much higher than today. Iron reduction is thought to be a major pathway for remineralization of organic carbon before the rise of oxygen (Konhauser et al. 2005, Posth et al. 2013), but if substantial iron

oxidation were not occurring (Box 1), there would be no significant source of iron oxides to support this respiratory process. Remineralization would instead be left with only sulfate reduction (which would be insignificant in the open ocean due low sulfate concentrations, though may have been more significant in locally sulfate-rich environments, Ueno 2014), fermentation, methanogenesis, and acetogenesis—metabolisms that are less energetic and less productive than respiratory pathways. Based on modern analog systems such as Lake Matano in Indonesia, the proportion of organic carbon that escapes remineralization and is buried in ferruginous systems is likely on the order of 20-25% of gross primary productivity (Crowe et al. 2011). Even higher burial fractions of 40-50% are expected for anoxic water columns based on extrapolation to zero of oxygen exposure time (a common metric for estimating the burial efficiency of an aquatic system) (Hartnett et al. 1998). In the absence of anaerobic electron acceptors like iron oxides, the fraction of organic carbon that is buried may have been even higher in the Archean. Assuming a burial fraction of ~50% as predicted for anoxic water columns (Hartnett et al. 1998), and rates of primary productivity predicted in Table 2, burial rates may have been ~1.37×10^{12} moles C/yr, only a few fold lower than modern marine carbon burial despite several orders of magnitude lower productivity.

Furthermore, reburial of eroded sedimentary carbon is a major flux of carbon on the Earth today, and could account for a large proportion of organic carbon in Archean sedimentary rocks, as kerogen would be less easily degraded under an anoxic atmosphere. In modern environments, erosion of sedimentary organic carbon leads to delivery to the ocean by rivers of about 8×10^{12} moles C/year (Galy et al. 2015, Daines et al. 2017). This

represents the fraction of eroded and resedimented fossil organic carbon that escapes oxygen; deriving the amount of fossil organic carbon that is initially eroded is challenging, as the fraction of eroded organic carbon that escapes oxidation is highly variable and depends on the maturity of the carbon and transport environment, but has been shown to vary within the range of about 10-80% (Bouchez et al. 2010, Hilton et al. 2014). Under an anoxic atmosphere, the amount of organic carbon reburied in sediments may have been much higher. However, even under the most conservative estimate based on modern environments, the resedimentation of eroded fossil organic carbon in the Archean was likely on the order of 10^{12} moles C/year, as much or even more than newly buried carbon.

Our understanding of the relative balances of carbon fixed and buried in Earth history are generally constrained by steady state models of the carbon cycle based on isotope mass balance of $\delta^{13}C$ in carbonates and organic carbon (e.g. Wickman 1956). These models are predicated on assumptions about steady state input and output of carbon through the fluid Earth, as well as the carbon isotope fractionations associated with carbon fixation. The value for fractionation of carbon isotopes during CO_2 fixation is generally assumed to be driven by RuBisCO (e.g. Berner 1991), the enzyme responsible for initial fixation of CO_2 in the Calvin cycle, used by all oxygenic phototrophs. Early in Earth history, however, this assumption may have been violated due to increased productivity from alternative carbon fixation pathways. Several alternative carbon fixation pathways, with fractionations both greater and lesser than that of the Calvin cycle, exist within bacteria and archaea (Hayes 2001). These alternative carbon fixation pathways are well established in niches focused around factors such as sensitivity to temperature and oxygen,

growth rate, and yield (Berg 2011). Based on these factors, the Calvin cycle is thought to be relatively young, specialized to tolerate oxygen (Berg 2011). The Wood-Ljungdahl (reverse acetyl-CoA) pathway is thought to be much more ancient, perhaps present in the Last Universal Common Ancestor (Lane et al. 2010). This pathway is characterized by very large fractionations between CO_2 and organic carbon (Hayes 2001), of up to -68 ‰ (Blaser et al. 2013).

The Archean carbonate carbon isotope record typically records values of ~0 ‰, similar to today, but the organic carbon isotope record is typically depleted relative to today (e.g. Fischer et al. 2009). This likely records increased contributions from alternative carbon fixation pathways with higher characteristic fractionations, including the reverse acetyl-CoA pathway used by methanogens and acetogens (Slotznick and Fischer 2016). As a result, similar carbonate carbon isotope values can be reached as today with a much lower fraction of carbon buried as organic matter (albeit with a more depleted carbon isotope value). For instance, given a $\delta^{13}C$ of carbonates of 0 ‰, the interpreted fraction of carbon buried as organic carbon would be ~0.25 assuming a fractionation characteristic of the Calvin cycle, but as low as ~0.09 given fractionations achieved by some organisms utilizing the reductive acetyl-CoA pathway. Assuming an input of carbon to the Earth system of similar magnitude to today, this lower fraction of carbon buried as organic carbon is in line with the burial rates estimated above, making the Archean carbon isotope record consistent with our estimates of productivity before the evolution of oxygenic photosynthesis.

Implications of electron limitation and excess nutrients in the pre-oxygen biosphere

If fluxes of nutrients like phosphate and fixed nitrogen were in excess of the demands of biological productivity before the evolution of oxygenic photosynthesis, this suggests that these nutrients may have accumulated to substantial concentrations in the oceans. This large reservoir of nutrients would have been immediately available to the first oxygenic phototrophs, fueling their almost unchecked growth for an initial, brief, intense period. This would be consistent with a rapid initial rise in oxygen concentrations following the evolution of oxygenic Cyanobacteria, leading to both a rapid Great Oxygenation Event and potentially an oxygen overshoot before depletion of this initial nutrient pool and development of a new steady state as suggested by a range of modeling and geochemical studies (e.g. Bekker and Holland 2012, Canfield et al. 2013, Ward et al. 2016, Luo et al. 2016). This is furthermore consistent with indications that a rapid accumulation of atmospheric oxygen and drawdown of CO_2 may have destabilized a methane greenhouse and triggered the Paleoproterozoic Snowball Earth Event (Kopp et al. 2005).

If nitrogen were not limiting to productivity before the rise of oxygen, there may not have been strong evolutionary pressure to evolve biological nitrogen fixation. Nitrogen fixation is a costly process, especially for the energy- and electron-limited organisms characteristic of the pre-oxygen biosphere, and so there would be a clear advantage to making use of abiotically fixed nitrogen rather than investing in fixing nitrogen via nitrogenase. While there may have still been locally nitrogen-limited environments conducive to the early evolution of nitrogenase, it is conceivable that the evolution of this enzyme (or at least emergence of the crown group forms that we see today), were relatively

late, postdating the evolution of anoxygenic photosynthesis, and potentially as late as the GOE. This is consistent with some phylogenetic and molecular clock analyses of the evolution of the nitrogenase family (e.g. Boyd et al. 2011) but will require further evidence to confidently assess.

Alternatively, nitrogenase may have evolved much earlier but for a different function and was only later coopted for fixation of N_2. Under high-methane conditions, the product of atmospheric (lightning and photochemical) nitrogen fixation would primarily be HCN (Navarro-Gonzalez et al. 2001, Tian et al. 2012). Following rainout to the ocean, HCN could hydrolyze to form NH_4OH (Zahnle 1986), but can also be taken up directly by biology via nitrogenase (Dekas et al. 2009). Large fluxes of HCN to the oceans may therefore have triggered the evolution of nitrogenase first as a way to detoxify and take up HCN, and was only later coopted to N_2 fixation, a scenario that has been hypothesized previously (e.g. Silver and Postgate 1973, Raymond 2005).

In the modern ocean, phosphate availability is often incredibly limited, leading to diverse strategies by phytoplankton to reduce their phosphate requirements such as substitution of membrane phospholipids with alternatives such as sulfonolipids (e.g. Van Mooy et al. 2006), sometimes pushing C:P ratios away from the Redfield value of 106:1 to values on the order of 500:1 (e.g. Bertilsson et al. 2003). However, these adaptations appear to be relatively recent, and phospholipids appear to be ancestral. If phosphate were much more abundant on the early Earth, this would explain the widespread use of phosphate head groups in membrane lipids, as they evolved when phosphate was available in great excess relative to productivity, and adaptations to low phosphate only evolved after

the evolution of oxygenic photosynthesis and drawdown of marine phosphate concentrations.

Our predictions of excess nutrient fluxes before the GOE have as corollaries testable hypotheses that can be verified with careful analysis of the rock record. One such testable hypothesis is that if phosphate fluxes were in excess of biological demand before the evolution of oxygenic photosynthesis, then the phosphorous cycle should have been dominated by abiotic processes, with the largest sinks for marine phosphorous being in the form of the precipitation and burial of phosphate minerals. If our model is correct, primary phosphate minerals—such as apatite—should have precipitated in the Archean ocean, potentially preserved in the rock record.

A major implication of this result is that without use of water or a similarly abundant electron donor, biology is fundamentally constrained in the mass and flux of biomass that can be fixed by the number of electrons supplied from geological processes. This was likely true on the early Earth, and would also be true of any biosphere on ice-covered worlds like Europa and Enceladus, where the use of water as an electron donor is impossible due to the lack of light penetration to the ocean. As a result, the size and productivity of any biosphere on these worlds would be limited by the flux of electron donor compounds from water/rock interactions at the base of their subsurface oceans. Furthermore, since anoxygenic photosynthesis would also be impossible in these environments, the efficiency and yield from consumption of these electron donors may be low, potentially more in line with Scenario 3 above than Scenarios 1 or 2. This would be true whether oxygen is supplied to the ocean from radiolysis of ice on the surface or not;

aerobic lithotrophic metabolisms typically have low efficiency per electron donor, of a similar order to methanogenesis (e.g. 40 Fe^{2+} oxidized to Fe^{3+} per 1 C fixed when using O_2, compared to 4-6 Fe^{2+} oxidized to Fe^{3+} per 1 C fixed when using light, Ehrenreich and Widdel 1994, Neubauer et al. 2002). This has substantial implications for the potential abundance, productivity, and detectability on icy moons, and should be kept in mind when planning life detection missions.

Table 1: Phosphate, fixed nitrogen, and electron donor fluxes to the pre-oxygen biosphere.

Flux	Nutrient flux (10^{12}moles/yr)	Citation	Notes
Phosphate (rivers)	0.0722	Tyrrel 1999	New nutrient
Phosphate (upwelling)	1.3	10% average modern marine deep water P concentrations per Bjerrum and Canfield 2002, textbook diffusivity for circulation rate	Includes recycling, high estimate using modern deep water phosphate concentration
Fixed N (photochemistry)	2	Tian et al. 2011	Production of HCN in relatively high CH_4 atmosphere
Fixed N (hydrothermal)	0.01	Brandes et al 1998	
Fixed N (lightning)	0.714	Navarro-Gonzalez et al. 1998, Navarro-Gonzalex et al. 2001	NO from high CO_2 atmosphere, or HCN from high CH_4
H_2 (volcanic outgassing)	4.7	Holland 2002	
H_2 (water-rock interactions)	1.36	Sherwood Lollar et al 2014	Sum of continental and oceanic fluxes
Fe^{2+} (hydrothermal)	0.19	Elderfield and Schultz 1996	
Fe^{2+} (production of new crust)	1.4	Lecuyer and Ricard 1999	

Table 2:

Productivity scenarios referenced in Figure 1. 1) Best estimate for electron donor-limited anoxygenic photosynthetic productivity. 2) Methanogen-dominated productivity. 3) Anoxygenic photosynthetic productivity without photosynthetic iron oxidation. 4) Modern marine oxygenic productivity.

Scenario	**Description**	**Primary Productivity (10^{12}moles C/yr)**	**Required P flux (10^{12}moles/yr)**	**Required N flux (10^{12}moles/yr)**
1	Electron donor limited anoxygenic photosynthesis (H_2 and Fe^{2+})	2.75	0.0259	0.415
2	Electron donor limited anoxygenic photosynthesis without photoferrotrophy (H_2)	2.43	.0229	0.367
3	Electron donor limited methanogenesis (H_2)	0.243	0.00229	0.0367
4	Modern marine productivity (oxygenic photosynthesis)	4000	37.7	604

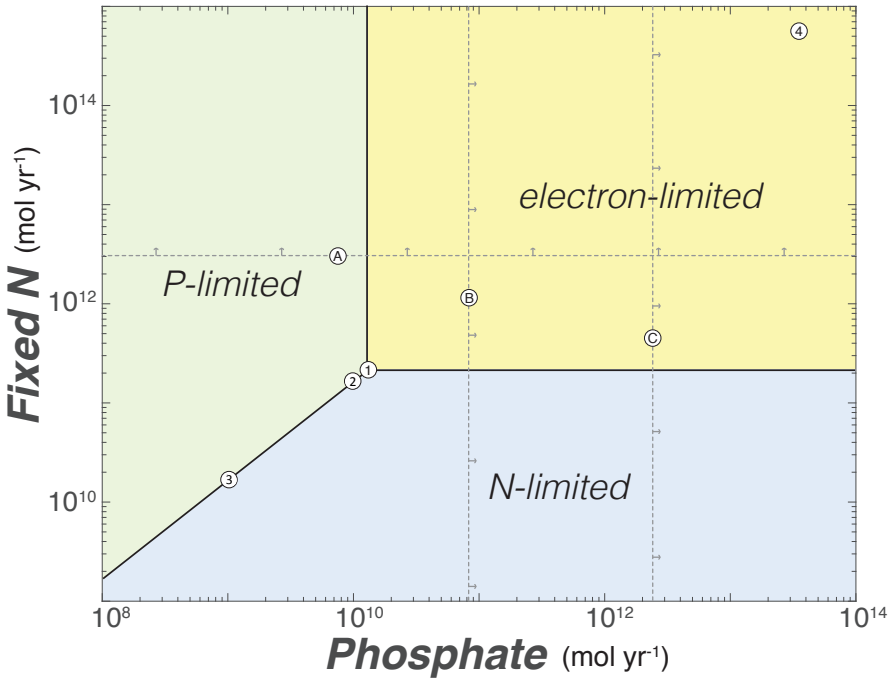


Figure 1:

Phosphate and fixed nitrogen fluxes to the biosphere, and resulting limitation on productivity. Electron donor, N, and P limitation fields assuming productivity in Scenario 1. Scenarios 2 and 3 shift the electron limited field down and to the left. Predicted nitrogen and phosphate fluxes (A-C) are well in excess of the demand of the electron limited rates of productivity predicted here.

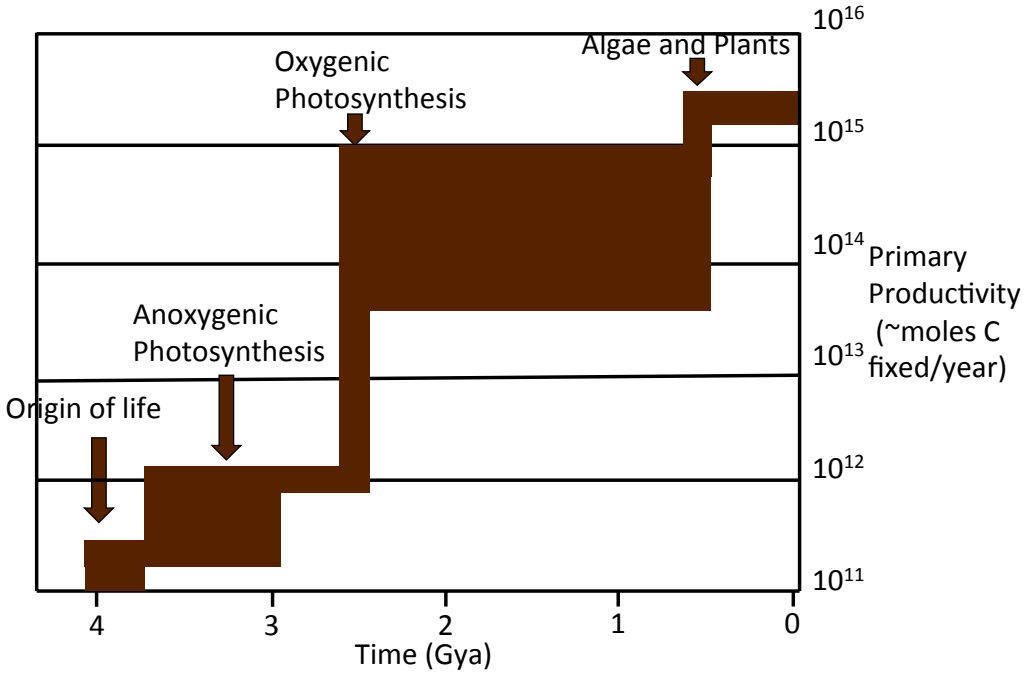


Figure 2:

Productivity of the biosphere (in terms of moles organic carbon fixed per year, on a log scale) through time (in billions of years). To first order, the productivity of life on Earth has increased through time, beginning near 0 at the origin of life to massive productivity today. The Great Oxygenation Event was associated with the largest boost to productivity since the origin of life, as a result of water splitting by oxygenic photosynthesis freeing life from electron donor limitation. An earlier, lesser jump to rates of productivity was likely associated with the evolution of anoxygenic photosynthesis, as a result of the conservation of life energy increasing the efficiency with which life could use electron donors to fix organic carbon.

Box 1: Did iron fuel the early biosphere?

Banded iron formations (BIFs) are among the most characteristic lithotypes in the Precambrian sedimentary record. These finely laminated, iron-rich (>15% Fe by weight) deposits are commonly thought to provide a record of biological activity on the early Earth (e.g. Harder 1919, Konhauser et al. 2002, Kappler et al. 2005). It is generally hypothesized that BIFs formed as a result of transport and concentration of ferrous iron (as $Fe^{2+}_{(aq)}$) in seawater under anoxic and sulfur-poor conditions, followed by oxidation and precipitation of iron as ferric iron phases (Holland 1973, Drever 1974, Holland 1984). The oxidation of iron in BIF deposition is commonly hypothesized to have been driven by biology, either indirectly by O_2 sourced from oxygenic photosynthesis by Cyanobacteria (e.g. Cloud 1973), or directly by aerobic iron-oxidizing bacteria (e.g. Chan et al. 2016) or anaerobic iron-oxidizing phototrophic bacteria (photoferrotrophs) (e.g. Widdel et al. 1993, Kappler et al. 2005). However, the complex depositional and alteration histories of ancient rocks have made the original composition of BIFs difficult to interpret, and as a result other hypotheses for pre-GOE BIF formation that do not rely on biological iron oxidation are viable. Demonstrating that early BIFs were not deposited via photoferrotrophy—and therefore not contributing to primary productivity—may have significant implications for our understanding of the Archean biosphere.

It is commonly thought that Archean BIFs were deposited as primary ferric iron oxides, and then reduced to ferrous minerals as a result of a combination of microbial iron

respiration and diagenesis (Konhauser et al. 2005, Fischer and Knoll 2009), and it is under this depositional model that biological iron oxidation mechanisms are invoked. However, a growing body of evidence suggests that Archean BIFs were deposited primarily as ferrous silicates (e.g. greenalite, Rasmussen et al. 2013, Rasmussen et al. 2015, Rasmussen et al. 2016; Rasmussen et al. 2017), a process that does not require microbial iron oxidation. In fact, if BIFs were being precipitated as ferrous minerals, this would not only be an absence of evidence for microbial iron oxidation, but would in fact be evidence of the absence of this process.

If Archean BIFs were precipitated as ferrous silicates, this suggests that photoferrotrophs were not oxidizing most of the iron in the Archean ocean. This raises an intriguing problem: if ferrous iron was abundant, and photoferrotrophy was a viable metabolism, why was it not occurring? Solutions to this problem could take one of a few forms. The first is that photoferrotrophy had not yet evolved. This seems an unlikely situation—photoferrotrophy seems to have been an ecologically viable niche in the Archean ocean, and is biochemically feasible given its presence in modern organisms (e.g. Widdel et al. 1993). For it not to have evolved by Archean time, as late as *ca* 2.5 Gya, when iron formations appear to have been deposited as ferrous silicates (e.g. Rasmussen et al. 2015, Rasmussen et al. 2016), runs counter to assumptions about the rapid evolution of microbes to exploit any available, thermodynamically viable redox couple (e.g. Broda 1977). It is especially curious given that photoferrotrophy appears to have evolved independently at least twice, in both the Chlorobi and the Proteobacteria (Ilbert and Bonnefoy 2013, Posth et al. 2013), implying that this metabolism is not as evolutionarily

difficult as other metabolisms such as oxygenic photosynthesis that have evolved only once (Falkowski 2011). A more compelling explanation is that photoferrotrophy was not a viable niche. This could be the case if, perhaps, ferrous iron were precipitated out of solution before reaching the photic zone. The flux of hydrothermal iron into an ocean with dissolved silica may have removed iron from the water column as amorphous ferrous silicate phases. In the modern ocean, dissolved silica concentrations are kept low due to the precipitation of silica biominerals, primarily by eukaryotic algae; in the Archean ocean, before the evolution of these oganisms, abiotic silica precipitation would have dominated, resulting in a more silica-saturated ocean (Stefurak et al. 2014). In an ocean where silica concentrations were close to saturation with respect to cristobalite or amorphous silica (Rasmussen et al. 2015), this process would titrate iron from the Archaean ocean, restricting the flux of ferrous iron to the photic zone, and limiting primary productivity by photoferrotrophy. Under this scenario, iron and light would not have substantially cooccurred in the Archean ocean, and photoferrotrophy would not have been a viable niche. This solution can be tested by investigating depth transects of BIF occurrences. Archean BIF are typically restricted to deep-water environments (Klein 2005), which is consistent with precipitation below the photic zone. In formations where depth transects can be traced between deep and shallow water facies, iron minerals are typically found only in deeper water facies (e.g. Tice and Lowe 2004). This could be consistent with dissolved iron not reaching the photic zone.

References:

1. Anbar, AD & Knoll, AH. Proterozoic ocean chemistry and evolution: a bioinorganic bridge? *Science* **297,** 1137–1142 (2002).
2. Battistuzzi FB, Feijão A, Hedges SB (2004) A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. BMC Evolutionary Biology 4, 44–57.
3. Bekker, A. & Holland, H. D. Oxygen overshoot and recovery during the early Paleoproterozoic. *Earth Planet. Sci. Lett.* **317–318,** 295–304 (2012).
4. Bekker, A. et al. Iron Formations: Their Origins and Implications for Ancient Seawater Chemistry. Treatise on Geochemistry: Second Edition 9, (2013).
5. Berg, I. a. Ecological aspects of the distribution of different autotrophic CO2 fixation pathways. *Appl. Environ. Microbiol.* **77,** 1925–1936 (2011).
6. Bertilsson, S., Berglund, O., Karl, D. M. & Chisholm, S. W. (2003) Limnol. Oceanogr. 48, 1721–1731.
7. Beukes, N. J. "Sedimentology of the Kuruman and Griquatown iron-formations, Transvaal Supergroup, Griqualand West, South Africa."Precambrian Research 24.1 (1984): 47-84.
8. Beukes NJ and Klein C (1990) Geochemistry and sedimentology of a facies transition— from microbanded to granular iron-formation—in the early Proterozoic Transvaal Supergroup, South Africa. Precambrian Research 47: 99–139.

9. Beukes, N.J., Klein, C., Kaufman, A.J., and Hayes, J.M., 1990, Carbonate petrography, kerogen distribution, and carbon and oxygen isotope variations in and early Proterozoic transition from limestone to iron forma- tion deposition, Transvaal Supergroup, South Africa: Economic Geology and the Bulletin of the Society of Economic Geologists, v. 85, p. 663–690.
10. Bird, L. J., Bonnefoy, V. & Newman, D. K. Bioenergetic challenges of microbial iron metabolisms. Trends Microbiol. 19, 330–340 (2011).
11. Blaser, M.B., Dreisbach, L.K., Conrad, R., 2013. Carbon isotope fractionation of 11 acetogenic strains grown on H2 and CO2. Appl. Environ. Microbiol. 79, 1787–1794.
12. Blättler, C. L. *et al*. Constraints on ocean carbonate chemistry and pCO2 in the Archaean and Palaeoproterozoic. *Nat*. *Geosci*. **1,** (2016).
13. Bontognali, T. R. R. *et al*. Sulfur isotopes of organic matter preserved in 3.45-billion-year-old stromatolites reveal microbial metabolism. *Proc*. *Natl*. *Acad*. *Sci*. **109,** 15146–15151 (2012).
14. Bontognali, T. R. R., Fischer, W. W. & Föllmi, K. B. Siliciclastic associated banded iron formation from the 3.2Ga Moodies Group, Barberton Greenstone Belt, South Africa. Precambrian Res. 226, 116–124 (2013).
15. Bouchez, J. *et al*. Oxidation of petrogenic organic carbon in the Amazon floodplain as a source of atmospheric CO2. *Geology* **38,** 255–258 (2010).
16. Boyd, E. S. *et al*. A late methanogen origin for molybdenum-dependent nitrogenase. *Geobiology* **9,** 221–232 (2011).

17. Boyd, E. S., Garcia Costas, A. M., Hamilton, T. L., Mus, F. & Peters, J. W. Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. *J. Bacteriol.* **197,** JB.02611–14 (2015).

18. Brandes, J. a *et al*. Abiotic nitrogen reduction on the early Earth. *Nature* **395,** 365–367 (1998).

19. Broda, E1. "Two kinds of lithotrophs missing in nature." *Zeitschrift für allgemeine Mikrobiologie* 17.6 (1977): 491-493.

20. Cairns-Smith, A.G., 1978, Precambrian solution photochemistry, inverse segregation, and banded iron formation: Nature, v. 276, p. 807–808.

21. Canfield, D. E., Rosing, M. T. & Bjerrum, C. Early anaerobic metabolisms. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361,** 1819-1834-1836 (2006).

22. Canfield, D. E., Glazer, A. N. & Falkowski, P. G. The evolution and future of Earth's nitrogen cycle. *Science* **330,** 192–196 (2010).

23. Canfield, D. E. *et al*. Oxygen dynamics in the aftermath of the Great Oxidation of Earth's atmosphere. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 16736–41 (2013).

24. Chan C. S., Fakra S. C., Emerson D., Fleming E. J., Edwards K. J. (2011). Lithotrophic iron-oxidizing bacteria produce organic stalks to control mineral growth: implications for biosignature formation. *ISME J* 5, 717–727.

25. Chan, C., Emerson, D. & Luther, G. W. The role of microaerophilic Fe-oxidizing microorganisms in producing banded iron formations. Geobiology (2016).

26. Crowe, S. A. *et al*. Photoferrotrophs thrive in an Archean Ocean analogue. *Proc. Natl. Acad. Sci.* **105,** 15938–15943 (2008).

27. Crowe, S. a. *et al*. The methane cycle in ferruginous Lake Matano. *Geobiology* **9,** 61–78 (2011).

28. Crowe, S. A. *et al*. Atmospheric oxygenation three billion years ago. *Nature* **501,** 535–538 (2013).

29. Crowe, S. a. *et al*. Sulfate was a trace constituent of Archean seawater. *Science (80-.)*. **346,** 735–739 (2014).

30. Daines, S. J., Mills, B. J. W. & Lenton, T. M. Atmospheric oxygen regulation at low Proterozoic levels by incomplete oxidative weathering of sedimentary organic carbon. *Nat. Commun.* **8,** 14379 (2017).

31. Drever, James I. "Geochemical model for the origin of Precambrian banded iron formations." Geological Society of America Bulletin 85.7 (1974): 1099-1106.

32. Ehrenreich, A., and Widdel, F., 1994, Anaerobic oxidation of ferrous iron by purple bacteria, a new type of phototrophic metabolism: Applied and Environmental Microbiology, v. 60, p. 4517–4526

33. Falkowski, P. G. The biological and geological contingencies for the rise of oxygen on earth. *Photosynth. Res.* **107,** 7–10 (2011).

34. Farquhar, J., Bao, H. & Thiemens, M. Atmospheric Influence of Earth's Earliest Sulfur Cycle. *Science (80-.)*. **289,** 756–758 (2000).

35. Fischer, W. W. & Knoll, A. H. An iron shuttle for deepwater silica in late Archean and early Paleoproterozoic iron formation. Bull. Geol. Soc. Am. 121, 222–235 (2009).

36. Fischer, W. W. et al. Isotopic constraints on the Late Archean carbon cycle from the Transvaal Supergroup along the western margin of the Kaapvaal Craton, South Africa. Precambrian Res. 169, 15–27 (2009).

37. Fischer, W., Hemp, J. & Johnson, J. E. Evolution of Oxygenic Photosynthesis. Annu. Rev. Earth Planet. Sci. 44, (2016).

38. Francois, L.M., 1986, Extensive deposition of banded iron formations was possible without photosynthesis: Nature, v. 320, p. 352–354.

39. Gauger, Tina, Kurt Konhauser, and Andreas Kappler. "Protection of phototrophic iron (II)-oxidizing bacteria from UV irradiation by biogenic iron (III) minerals: Implications for early Archean banded iron formation." *Geology*43.12 (2015): 1067-1070.

40. Goldblatt, C., Lenton, T. M. & Watson, A. J. Bistability of atmospheric oxygen and the Great Oxidation. *Nature* **443,** 683–686 (2006).

41. Habicht, K. S., Gade, M., Thamdrup, B., Berg, P. & Canfield, D. E. Calibration of sulfate levels in the archean ocean. *Science* **298,** 2372–2374 (2002).

42. Hartnett, H., Keil, R., Hedges, J. & Devol, A. Influence of oxygen exposure time on organic carbon preservation in continental margin sediments. *Nature* **391,** 572–575 (1998).

43. Hayes, J. M. Fractionation of the Isotopes of Carbon and Hydrogen Processes Biosynthetic Processes. *Rev. Mineral. Geochemistry* **43,** 225–277 (2001).

44. Heising, Silke, et al. "Chlorobium ferrooxidans sp. nov., a phototrophic green sulfur bacterium that oxidizes ferrous iron in coculture with a “Geospirillum” sp. strain." *Archives of Microbiology* 172.2 (1999): 116-124.

45. Higgins, J. A., Fischer, W. W. & Schrag, D. P. Oxygenation of the ocean and sediments : Consequences for the sea fl oor carbonate factory. *Earth Planet. Sci. Lett.* **284,** 25–33 (2009).

46. Hilton, R. G., Gaillardet, J., Calmels, D. & Birck, J. L. Geological respiration of a mountain belt revealed by the trace element rhenium. *Earth Planet. Sci. Lett.* **403,** 27–36 (2014).

47. Holland, Heinrich D. "The oceans; a possible source of iron in iron-formations." Economic Geology 68.7 (1973): 1169-1172.

48. Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press

49. Holland, H. D. Volcanic gases, black smokers, and the great oxidation event. *Geochim. Cosmochim. Acta* **66,** 3811–3826 (2002).

50. Johnson, J. E. et al. Manganese-oxidizing photosynthesis before the rise of cyanobacteria. Proc. Natl. Acad. Sci. 110, 11238–11243 (2013a).

51. Johnson, J. E. et al. Reply to Jones and Crowe: Correcting mistaken views of sedimentary geology, Mn-oxidation rates, and molecular clocks. Proc. Natl. Acad. Sci. 110, E4119–E4120 (2013b).

52. Johnson, J. E., Gerpheide, A., Lamb, M. P. & Fischer, W. W. O2 constraints from Paleoproterozoic detrital pyrite and uraninite. *Geol. Soc. Am. Bull.* **126,** 813–830 (2014).

53. Kappler, A. & Newman, D. K. Formation of Fe(III)-minerals by Fe(II)-oxidizing photoautotrophic bacteria. Geochim. Cosmochim. Acta 68, 1217–1226 (2004).

54. Kappler, A., Pasquero, C., Konhauser, K. O. & Newman, D. K. Deposition of banded iron formations by anoxygenic phototrophic Fe(II)-oxidizing bacteria. Geology 33, 864–865 (2005).

55. Kasting, J. F. Theoretical constraints on oxygen and carbon dioxide concentrations in the Precambrian atmosphere. *Precambrian Res.* **34,** 205–229 (1987).

56. Kasting, J. F. Earth's Early Atmosphere. *Science (80-.).* **259,** 920–926 (1993).

57. Kasting, J. F., Pavlov, A. A. & Siefert, J. L. A coupled ecosystem-climate model for predicting the methane concentration in the Archean atmosphere. *Orig. Life Evol. Biosph.* **31,** 271–285 (2001).

58. Kharecha, P., Kasting, J. & Siefert, J. A coupled atmosphere–ecosystem model of the early Archean Earth. *Geobiology* **3,** 53–76 (2005).

59. Klein, C. Some Precambrian banded iron-formations (BIFs) from around the world: Their age, geologic setting, mineralogy, metamorphism, geochemistry, and origin. Am. Mineral. 90, 1473–1499 (2005).

60. Konhauser, K. O. et al. Could bacteria have formed the Precambrian banded iron formations? Geology 30, 1079–1082 (2002).

61. Konhauser, K. O., Newman, D. K. & Kappler, A. The potential significance of microbial Fe (III) reduction during deposition of Precambrian banded iron formations. Geobiology 3, 167–177 (2005).

62. Kopp, R. E., Kirschvink, J. L., Hilburn, I. A. & Nash, C. Z. The Paleoproterozoic snowball Earth: a climate disaster triggered by the evolution of oxygenic photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 11131–11136 (2005).

63. Lane, N., Allen, J. F. & Martin, W. How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays* **32,** 271–280 (2010).

64. Leigh, John A. "Nitrogen fixation in methanogens: the archaeal perspective." *Current issues in molecular biology* 2 (2000): 125-131.

65. Llirós, M. *et al*. Pelagic photoferrotrophy and iron cycling in a modern ferruginous basin. *Sci. Rep.* 1–5 (2015). doi:10.1007/s13398-014-0173-7.2

66. Luo, G. *et al*. Rapid oxygenation of Earths atmosphere 2.33 billion years ago. *Sci. Adv.* **2,** e1600134–e1600134 (2016).

67. Lyons, T. W., Reinhard, C. T. & Planavsky, N. J. The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506,** 307–15 (2014).

68. Navarro-González, R., Molina, M. J. & Molina, L. T. Nitrogen fixation by volcanic lightning in the early Earth. *Geophys. Res. Lett.* **25,** 3123 (1998).

69. Navarro-gonzalez, R. *et al*. A possible nitrogen crisis for Archaean life due to reduced nitrogen fixation by lightning. *Nature* **412,** 61–64 (2001).

70. Neubauer, S. C. et al. Life at the energetic edge: kinetics of circumneutral Fe oxidation by lithotrophic iron oxidizing bacteria isolated from the wetland plant rhizosphere. Appl. Environ. Microbiol. 68, 3988–3995 (2002).

71. Olson, J. M. 2006 Photosynthesis in the Archean era. Photosynth. Res. 88, 109–117. (doi:10.1007/s11120-006-

72. Ono S, Eigenbrode JL, Pavlov AA, Kharecha P, Rumble DI, Kasting JF, Freeman KH (2003) New insights into Archean sulfur cycle from mass-independent sulfur isotope records. Earth and Planetary Science Letters 213, 15–30.

73. Pancost, R. D. & Sinninghe Damsté, J. S. Carbon isotopic compositions of prokaryotic lipids as tracers of carbon cycling in diverse settings. *Chem. Geol.* **195,** 29–58 (2003).

74. Pavlov, A., Kasting, F., Brown, L. L., Rages, K. a & Freedman, R. Greenhouse warming by CH4 in the atmosphere of early Earth. *J. Geophys. Res.* **105,** 11981–11990 (2000).

75. Pavlov, A. A., Kasting, J. F., Eigenbrode, J. L. & Freeman, K. H. Organic haze in Earth's early atmosphere: Source of low-13C Late Archean kerogens? *Geology* **29,** 1003–1006 (2001).

76. Planavsky, N. J. *et al*. Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. *Nat. Geosci.* **7,** 283–286 (2014).

77. Posth, N. R., Konhauser, K. O. & Kappler, A. Microbiological processes in banded iron formation deposition. Sedimentology 60, 1733–1754 (2013).

78. Rasmussen, B., Meier, D. B., Krapež, B. & Muhling, J. R. Iron silicate microgranules as precursor sediments to 2.5-billion-yearold banded iron formations. Geology 41, 435–438 (2013).

79. Rasmussen, B., Krapež, B., Muhling, J. R. & Suvorova, A. Precipitation of iron silicate nanoparticles in early Precambrian oceans marks Earth's first iron age. Geology 43, 303–306 (2015).

80. Rasmussen, B., Muhling, J. R., Suvorova, A. & Krapež, B. Dust to dust: Evidence for the formation of 'primary' hematite dust in banded iron formations via oxidation of iron silicate nanoparticles. Precambrian Res. 284, 49–63 (2016).

81. Rasmussen, B., Muhling, J. R., Suvorova, A., and Krapez, B., 2017, Greenalite precipitation linked to the deposition of banded iron formations downslope from a late Archaean carbonate platform. *Precambrian Research*, **290**, 49-62.

82. Raymond, J., Siefert, J. L., Staples, C. R. & Blankenship, R. E. The Natural History of Nitrogen Fixation. *Mol. Biol. Evol.* **21,** 541–554 (2004).

83. Robbins, L. J. *et al*. Trace elements at the intersection of marine biological and geochemical evolution. *Earth-Science Rev.* **163,** 323–348 (2016).

84. Rosing, M. T. 13C-Depleted Carbon Microparticles in >3700-Ma Sea-Floor Sedimentary Rocks from West Greenland. *Science (80-.).* **283,** 674–676 (1999).

85. Rye, R., Kuo, P. H. & Holland, H. D. Atmospheric carbon dioxide concentrations before 2.2 billion years ago. *Nature* **378,** 603–605 (1995).

86. Sagan, C. & Mullen, G. Earth and Mars: Evolution of Atmospheres and Surface Temperatures. *Science (80-.).* **177,** 52–56 (1972).

87. Saito, M. A., Sigman, D. M. & Morel, F. M. M. The bioinorganic chemistry of the ancient ocean: the co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean–Proterozoic boundary? *Inorganica Chim. Acta* **356,** 308–318 (2003).

88. Schidlowski, M., Appel, P. W. ., Eichmann, R. & Junge, C. E. Carbon isotope geochemistry of the 3.7 × 109-yr-old Isua sediments, West Greenland: implications for the Archaean carbon and oxygen cycles. *Geochim. Cosmochim. Acta* **43,** 189–199 (1979).

89. Sherwood Lollar, B., Onstott, T. C., Lacrampe-Couloume, G. & Ballentine, C. J. The contribution of the Precambrian continental lithosphere to global H2 production. *Nature* **516,** 379–82 (2014).

90. Shih, P., Hemp, J., Ward, L., Matzke, N. & Fischer, W. 2016. Crown group oxyphotobacteria postdate the rise of oxygen. Geobiology.

91. Silver, W. S., and J. R. Postgate. "Evolution of asymbiotic nitrogen fixation." *Journal of Theoretical Biology* 40.1 (1973): 1-10.

92. Slotznick, S. P. & Fischer, W. W. Examining Archean methanotrophy. *Earth Planet. Sci. Lett.* **441,** 52–59 (2016).

93. Stefurak, Elizabeth JT, et al. "Primary silica granules—A new mode of Paleoarchean sedimentation." *Geology* 42.4 (2014): 283-286.

94. Stüeken, E. E., Buick, R., Guy, B. M. & Koehler, M. C. Isotopic evidence for biological nitrogen fixation by molybdenum-nitrogenase from 3.2 Gyr. *Nature* 1–9 (2015). doi:10.1038/nature14180

95. Tian, F., Kasting, J. F. & Zahnle, K. Revisiting HCN formation in Earth's early atmosphere. *Earth Planet. Sci. Lett.* **308,** 417–423 (2011).

96. Tice, M. M. & Lowe, D. R. Photosynthetic microbial mats in the 3 , 416-Myr-old ocean. *Nature* **431,** 549–552 (2004).

97. Tice, M. M. & Lowe, D. R. Hydrogen-based carbon fixation in the earliest known photosynthetic organisms. *Geology* **34,** 34–37 (2006).

98. Tosca, Nicholas J., Stephen Guggenheim, and Peir K. Pufahl. "An authigenic origin for Precambrian greenalite: Implications for iron formation and the chemistry of ancient seawater." Geological Society of America Bulletin 128.3-4 (2016): 511-530.

99. Ueno, Yuichiro. "Coping with low ocean sulfate." *science* 346.6210 (2014): 703-704.

100. Walker, J. C. & Brimblecombe, P. Iron and sulfur in the pre-biologic ocean. *Precambrian Res.* **28,** 205–222 (1985).

101. Ward, L. M., Kirschvink, J. L. & Fischer, W. W. 2016. Timescales of Oxygenation Following the Evolution of Oxygenic Photosynthesis. Orig. Life Evol. Biosph. 46(1) pp51-65.

102. Ward, LM, A Idei, T Kakegawa, WW Fischer, and SE McGlynn. Microbial diversity and iron oxidation at Okuoku-hachikurou Onsen, a Japanese hot spring analog of Precambrian iron formation. Geobiology, submitted.

103. Weiss, M. C. *et al.* The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* 1–8 (2016). doi:10.1038/nmicrobiol.2016.116

104. Wheat, C. Geoffrey, Richard A. Feely, and Michael J. Mottl. "Phosphate removal by oceanic hydrothermal processes: An update of the phosphorus budget in the oceans." *Geochimica et Cosmochimica Acta* 60.19 (1996): 3593-3608.

105. Wickman, Frans E. "The cycle of carbon and the stable carbon isotopes." *Geochimica et cosmochimica acta* 9.3 (1956): 136-153.

106. Widdel, F., Schnell, S., Heising, S., Ehrenreich, A., Assmus, B., and Schink, B., 1993, Ferrous iron oxidation by anoxygenic phototrophic bacteria: Nature, v. 362, p. 834–836.

107. Williams, Robert Joseph Paton, and Ros Rickaby. *Evolution's destiny: co-evolving chemistry of the environment and life*. Royal Society of Chemistry, 2012.

108. Zahnle, K. J. Photochemistry of methane and the formation of hydrocyanic acid (HCN) in the Earth's early atmosphere. *J. Geophys. Res.* **91,** 2819–2834 (1986).

109. Zerkle, A. L. & Mikhail, S. The geobiological nitrogen cycle: From microbes to the mantle. Geobiology 1–10 (2017). doi:10.1111/gbi.12228

Chapter 2

TIMESCALES OF OXYGENATION FOLLOWING THE EVOLUTION OF OXYGENIC PHOTOSYNTHESIS

Ward, Lewis M., Joseph L. Kirschvink, and Woodward W. Fischer. "Timescales of oxygenation following the evolution of oxygenic photosynthesis." Origins of Life and Evolution of Biospheres 46.1 (2016): 51-65. DOI: 10.1007/s11084-015-9460-3

Abstract

Among the most important bioenergetic innovations in the history of life was the invention of oxygenic photosynthesis—autotrophic growth by splitting water with sunlight—by Cyanobacteria. It is widely accepted that the invention of oxygenic photosynthesis ultimately resulted in the rise of oxygen by *ca*. 2.35 Gya, but it is debated whether this occurred more or less immediately as a proximal result of the evolution of oxygenic Cyanobacteria or whether they originated several hundred million to more than one billion years earlier in Earth history. The latter hypothesis involves a prolonged period during which oxygen production rates were insufficient to oxidize the atmosphere, potentially due to redox buffering by reduced species such as higher concentrations of ferrous iron in seawater. To examine the characteristic timescales for environmental oxygenation following the evolution of oxygenic photosynthesis, we applied a simple mathematical approach that captures many of the salient features of the major biogeochemical fluxes and reservoirs present in Archean and early Paleoproterozoic surface environments. Calculations illustrate that oxygenation would have overwhelmed redox buffers within

~100 kyr following the emergence of oxygenic photosynthesis, a geologically short amount of time unless rates of primary production were far lower than commonly expected. Fundamentally, this result arises because of the multiscale nature of the carbon and oxygen cycles: rates of gross primary production are orders of magnitude too fast for oxygen to be masked by Earth's geological buffers, and can only be effectively matched by respiration at non-negligible O_2 concentrations. These results suggest that oxygenic photosynthesis arose shortly before the rise of oxygen, not hundreds of millions of years before it.

Introduction:

Sometime between 2.4 and 2.35 billion years ago (Gya) the Earth experienced one of its largest and most significant changes, when free molecular oxygen first accumulated in the atmosphere (Bekker et al. 2004, Papineau et al. 2007, Guo et al. 2009, Hoffman 2013, Johnson et al. 2013, Johnson et al. 2014, Rasmussen et al. 2013). This event has long been recognized from a wide range of geological and geochemical proxies (Figure 1), and is thought to be caused by the metabolic products of oxygenic Cyanobacteria (e.g. Falkowski 2011, Shih 2015). However, there is a long history of interpretations and hypotheses regarding the greater antiquity of Cyanobacteria deep into Archean time that have proven controversial. Interpretations in support of Cyanobacteria in Paleoarchean time once included putative cyanobacterial microfossils from the Archean Apex Chert (Schopf 1993), which remain uncertain as body fossils regardless of phylogenetic affinity (Brasier et al. 2002). Even if these structures are proven to be microfossils, convergence of morphology makes the interpretation of these and other filamentous microfossils as

Cyanobacteria equivocal (Knoll and Golubic 1992, Shih et al. 2013). Similarly, stromatolites—accretionary sedimentary growth structures often interpreted to be formed by the interaction of microbial mats and sediment—have been interpreted as cyanobacterial products in deep time. The stromatolite record dates back more than 3.4 Ga (Hofmann et al. 1999, Allwood et al. 2006), with many stromatolites interpreted as being formed by Cyanobacteria (e.g. Buick 1992). However, it has been proposed that ancient stromatolites need not have been formed by Cyanobacteria (Brock 1978, Bosak et al. 2007), and hypothesized that not all stromatolites are formed biogenically (Grotzinger and Knoll 1999), making the relationship between them and Cyanobacteria uncertain. Schidlowski and colleagues (1975, 1979) argued on the basis of carbonate carbon isotope values that oxygenic photosynthesis had been in place since at least 3.7 Gya, with Rosing and Frei (2002) reaching the same conclusion from observations of ^{13}C-depleted graphite and relatively high lead concentrations in graphite-bearing quartz ± garnet schists of similar age; however, none of these features may be diagnostic of oxygenic photosynthesis (Kopp et al. 2005). The discovery of 2-methylhopane biomarkers in Archean rocks has also been used to argue for Cyanobacteria predating the rise of oxygen (Brocks et al. 1999), although from subsequent studies it has been recognized that these molecules are not unique to Cyanobacteria (Rashby et al. 2007, Welander et al. 2010), likely evolved in other phyla (Ricci et al. 2015), and may not even be native to the rocks (French et al. 2015). Most recently, a range of data from trace metal proxies have been interpreted as "whiffs" of oxygen in Archean surface environments, more specifically spatially or temporally local pulses of cyanobacterial O_2 that left a record in trace metal oxygen proxies but failed to

fully or irreversibly oxidize the atmosphere (Anbar et al. 2007, Kaufman et al. 2007). Recent trace metal data from strata nearly three billion years old have been interpreted to show evidence of oxygenic photosynthesis (Planavsky et al. 2014, Crowe et al. 2013). The interpretation of these trace metal signatures of O_2 remains controversial in part because their geochemical cycles are not well understood both in modern environments and in diagenetically stabilized and post-depositionally altered lithologies typical of Precambrian successions (Helz et al. 2011, Nägler et al. 2011, Morford et al. 2012); these whiff signatures also appear to conflict with independent geochemical O_2 proxies such as redox-sensitive detrital grains and mass independent fractionation of sulfur isotopes (e.g. Johnson et al. 2014). Furthermore, the long stretch of geological time over which whiffs are thought to have occurred is at odds with expectations for the productivity of oxygenic photosynthesis, and so limits to the spread of oxygen are invoked (e.g. Lyons et al. 2014). These arguments often rely on redox buffers, geologically sourced reduced compounds in the atmosphere and/or oceans (like Fe^{2+} or CH_4), which reacted with molecular oxygen to prevent its environmental accumulation (Schidlowski 1983, Gaillard et al. 2011, Kump and Barley 2007). The depletion of these redox buffers over geological time due either to changes in source fluxes or reaction with oxygen is thought to eventually allow oxygen to rise (e.g. Holland 2009). The "rusting" of the fluid Earth has long been considered as a mechanism for the deposition of banded iron formations (Cloud 1973, Walker et al. 1983), though these deposits are now recognized to have much more complex origins and are comosed of dominantly ferrous mineral products from an iron cycle that need not have involved molecular oxygen (Fischer and Knoll 2009, Rasmussen et al. 2013b). Ultimately,

however, it remains unclear how well geological processes and redox buffers might counter the large fluxes and rapid responses anticipated of biological productivity.

Massive numbers characterize the modern oxygen cycle, including fluxes that appear capable of responses to perturbations on extremely short timescales. The atmosphere currently contains nearly 21% dioxygen by volume (~3.8x 10^{19} moles O_2), and this concentration is thought to have been largely stable at least since the beginning of the Cenozoic Era (Glasspool and Scott 2010). Global annual net primary productivity (NPP) is estimated from measurements of carbon fluxes to be on the order of 105 Pg C/year, an amount equivalent to ~8.75 x 10^{16} moles of O_2 per year (Field et al. 1998), giving a residence time of O_2 in the atmosphere of only about 4300 years. Such a short residence time of atmospheric O_2 is further supported by O isotope ratio data on O_2 from gas trapped within ice cores, which shows large and rapid variation over geologically short (< 1 kyr) timescales, implying a residence time for O_2 in the atmosphere perhaps as short as a couple of hundred years (Severinghaus et al. 2009). Thus any accurate first order description of the oxygen cycle should capture these rapid dynamics. Indeed it is remarkable that atmospheric oxygen appears to remain stable over a range of timescales (e.g. Suwa and Bender 2008, Berner 1989) despite rapid turnover and variability in global primary productivity over glacial/interglacial timescales (Bender et al. 1994, Paytan et al. 1996) to Phanerozoic timescales (Robinson 1990). Tight feedbacks must be in place to maintain oxygen concentrations.

The largest sink in the oxygen cycle is that of aerobic respiration. While gross primary productivity (GPP, the rate at which organic carbon is fixed by organisms) is very

large, about half of all fixed carbon is respired by plants and algae themselves (Bender et al. 1999). The remaining fixed carbon is referred to as net primary productivity (NPP), and the vast majority of this is remineralized by heterotrophs. Only a small fraction is ultimately buried and allows the accumulation of O_2 in the atmosphere, on the order of 0.32% of NPP (Berner 1989). This tight coupling of primary production and aerobic respiration has maintained a stably oxygenated atmosphere over Phanerozoic time (Glasspool and Scott 2010), but this balance may not have applied soon after oxygenic photosynthesis evolved, as efficient mechanisms for aerobic respiration may not have evolved until after the evolution of oxygenic photosynthesis (Gribaldo et al. 2009). Furthermore, aerobic respiration typically requires a threshold concentration of O_2 in the environment to be efficient, with most organisms transitioning to anaerobic metabolism at oxygen concentrations lower than ~3 μM, near what is known as the Pasteur Point (Canfield 2005). Slow growth of a strain of *Escherichia coli* was observed to occur at lower, nanomolar oxygen concentrations (Stolper et al. 2010), though the ecological and environmental relevance of these results has been questioned (Ducluzeau et al. 2014). The evolutionary timing and oxygen requirements of aerobic respiration remain important open questions, but it is reasonable to infer that early respiration was slow due to oxygen concentrations far below the concentrations at which respiratory enzymes optimally function (e.g. Guo et al. 2009). The largest control on preservation of organic carbon in the sediments is oxygen exposure time (Hartnett et al. 1998). As oxygen exposure time would be much lower before and during the rise of oxygen, organic carbon burial efficiency would almost certainly be higher than today. This is supported by measurements of organic

carbon burial rates in Lake Matano, an interesting Archean analog site, where burial rates are on the order of 22% of NPP (Crowe et al. 2011). Lake Matano is a permanently stratified lake in Indonesia and is considered one of the best extant Archean ocean analog systems due to its photic zone anoxia, high dissolved iron content, and low sulfur content (Crowe et al. 2008, Crowe et al. 2011, Jones et al. 2011). The absence of aerobic respiration before the rise of oxygen likely required alternative remineralization processes to play larger roles in the carbon cycle, most likely dominated by methanogenesis (Hayes 1994). This has been observed in Lake Matano, where methanogenesis appears to be responsible for ~50% of organic carbon remineralization (Crowe et al. 2011).

Until oxygen concentrations were reached that made aerobic respiration an efficient O_2 sink, O_2 sourced from oxygenic phototrophs would largely titrate reduced compounds such as methane and ferrous iron in seawater and the atmosphere—redox buffers that might prevent accumulation of oxygen in the atmosphere depending on their relative abundances and reaction kinetics (Lyons et al. 2014). It is critical to note here that many of these redox buffers are ultimately replenished on geological (i.e. chemical weathering) timescales—these are slow relative to phototrophic oxygen fluxes. This mismatch of timescales suggests the hypothesis that, because redox buffers would have been replenished sufficiently slowly, O_2 would have accumulated rapidly from the perspective of geological time. To test this idea and quantitatively determine whether these limitations would be sufficient to slow the oxygenation of the atmosphere over the long term, we present a geochemical box model of the Archean Earth that includes a realistic assessment of the oxygen sinks possibly present in the environment. Rather than hundreds of millions of years, we find that oxygenation

should have happened in tens of thousands to perhaps a hundred thousand years after oxygenic photosynthesis evolved in Cyanobacteria.

Basic accounting

An initial, simple estimate of the time necessary to oxygenate the atmosphere can be determined simply by summing the pools of reduced "redox buffer" compounds in the atmosphere and ocean and dividing these by the gross O_2 flux. Using estimates of reduced compound concentrations and rates of productivity and carbon burial detailed below, this results in the consumption of all stored reducing power on Earth's surface within a few hundred years. While this serves as no more than a rough approximation, it highlights that despite the overall reduced state of the Archean Earth, absolute abundances of redox buffer compounds are small compared to the anticipated fluxes of oxygen produced through oxygenic photosynthesis. A more sophisticated oxygenation model, accounting for geological and biological fluxes of reductants into the Earth system, is detailed below.

Model summary

We follow the estimates detailed below for Archean starting concentrations and fluxes of each oxygen sink. However, we make a number of simplifying assumptions, including: (1) the oceans are of the same volume as the modern (~1.3 x 10^{21} liters), (2) the atmosphere was made up of the same number of gas molecules as today (~1.8 x 10^{20} moles), and (3) primary productivity driven by anoxygenic photosynthesis is minor relative to oxygenic photosynthesis (e.g. Kharecha et al. 2005). Furthermore, (4) we note that the net reactions

that describe anoxygenic photosynthesis are equivalent to oxygenic photosynthesis when O_2 equivalents are measured (e.g. iron oxides are produced equivalently by direct photosynthetic oxidation or via oxidation by O_2 produced from oxygenic photosynthesis). This allows us to consider primary productivity as a single flux. Additionally, (5) we ignore sulfide fluxes because any such input into the fluid Earth will be removed as pyrite as it reacts with ferrous iron in the predominantly ferruginous oceans. This follows logic from Walker and Brimblecombe (1995), whereby larger geochemical iron sources will lead to removal of sulfur unless the iron is first immobilized by oxidation. As this will also serve as a sink for reduced iron, ignoring sulfide can be considered a conservative estimate. Next, (6) the ocean is assumed to be well-mixed with regards to chemical compounds and microbial populations. While ocean and atmospheric circulation are important factors in controlling the distribution of microbes and geochemical compounds, these processes are relatively rapid, on the order of ~1 kyr—a timescale much finer than the typical resolution of the sedimentary record (Sadler 1981). While the oceans likely cannot stagnate completely, they can be dynamically stratified as a result of changes in circulation patterns (e.g. Hotinski et al. 2000, Hotinski et al. 2001), but even if mixing were substantially slower this would only serve to isolate reduced compounds to the deep ocean and would allow the surface oceans and the atmosphere to oxidize even faster. Similarly, oxidation rates of all species by O_2 are considered to be geologically instantaneous, so kinetic reaction rates are ignored. However, rates are included in the particular cases of hydrogen escape and methane photolysis, as these are driven by processes other than the production of O_2.

Reduced pools. A number of pools of reduced compounds are expected to be present in the ocean and atmosphere before the rise of oxygen, and are summarized in Figure 2. Major atmospheric reduced species include hydrogen and methane. Estimates of atmospheric mixing ratios of 0.001 (i.e. 1000 ppmv, or 1 permil) for these species were derived from Pavlov et al. (2000) [later estimates by Haqq-Misra et al (2011) revised estimates for Archean H_2 mixing ratio down to 0.0003, but we use here the earlier, higher value as a conservative estimate]. These represent upper bounds for the concentrations of each of these species in order to satisfy the Faint Young Sun paradox (Pavlov et al. 2000). If the Faint Young Sun is instead counteracted by, for instance, high concentrations of CO_2 as has been elsewhere proposed (Owan et al. 1979), these concentrations may be substantially lower and would lower existing reduced pools and accelerate atmospheric oxidation; we therefore utilize these upper concentration ranges as a conservative estimate.

Major aquatic reduced pools are dissolved Fe^{2+}, Mn^{2+}, and ammonia. Estimates were derived from Holland (1984) and Canfield et al. (2010), and range up to 120 micromolar for iron and manganese, and 80 micromolar for ammonia. Metal concentration estimates come from the solubility of their respective carbonate species, while ammonium concentrations were based on an assumption of nutrient co-limitation with iron by anoxygenic phototrophs prior to the advent of oxygenic photosynthesis. Large ranges of possible concentrations were proposed; the highest estimates were used here.

Reduced fluxes. Fluxes of reduced compounds to the atmosphere and oceans from processes such as hydrothermal cycling and volcanic outgassing are summarized in Figure 3. Fluxes of reduced compounds from volcanoes were estimated based on Holland (2002) to be a combined 1.4 x 10^{12} moles of SO_2, CO, and H_2 with identical oxidation stoichiometries. If H_2 is a significant volcanic gas, this could increase overall rates of hydrogen escape, driving an O_2-independent sink of reducing power, increasing the rate of atmospheric oxidation. Pooling all volcanic gases therefore provides a conservative estimate.

Production of oceanic crust and associated reduced iron was derived from Lécuyer and Ricard (1999). Based on Williams et al. (2012), it is assumed that mantle redox—and therefore the Fe^{2+}/ Fe^{3+} ratio of new crustal material—has been constant since accretion.

Hydrothermal fluxes were derived from Elderfield and Schultz (1996) for modern values as a reference, and Archean fluxes are naively assumed to be ten times modern based on Des Marais (2000), although it is possible that Archean hydrothermal fluxes were lower than modern (e.g. Korenaga 2006). While Fe^{2+} is here given as a major component of hydrothermal fluxes, this iron is sourced from weathering of newly formed crust. Because it is tricky to untangle this, this iron is therefore essentially double-counted in our model here to provide a conservative oxygenation timescale estimate.

Biogenic methane fluxes are largely unconstrained for Archean time, but solutions for the Faint Young Sun paradox using methane as a greenhouse gas assume comparable fluxes to the modern (e.g. Pavlov et al. 2001)—an estimate arrived at independently by Kharecha et al. (2005). This is equivalent to about 8% of remineralized organic matter

being converted to methane given the rates of primary productivity used in this model. Estimates for methanogenic rates can also be made by analogy to Lake Matano (Crowe et al. 2008). It has been demonstrated that in Lake Matano, 50% or more of fixed organic carbon is remineralized via methanogenesis (Crowe et al. 2011). This is consistent with estimates from Hayes (1994), who hypothesized from carbon isotope mass balance that CO_2 and CH_4 were equally significant for Archean carbon cycling.

Oxidizing fluxes. Photolysis of methane, and hydrogen escape to space are modeled as first order reactions with time constants based on rates from Catling et al. (2001), resulting in a net loss of reducing power from the Earth system over time without consumption of O_2. Methane photolysis in the Archean atmosphere is expected to be limited by concentrations of hydroxyl radicals derived from H_2O photolysis, and so is expected to be ~1000 times slower than modern methane breakdown in an oxygen-rich atmosphere (Pavlov et al. 2000). The flux of fixed carbon to methane is a significant variable in this model. As methane can be lost from the Earth system through an O_2-independent photolytic reaction, it acts as an absolute sink of electron equivalents, allowing oxygen to react with other reductants and eventually to build up. We have therefore used a more conservative, low value for methanogenic flux.

While the nutrient status of Archean seawater is largely unknown, we take primary productivity as limited to 10% of modern marine values due to phosphorous limitation as suggested by Bjerrum and Canfield (2002), resulting in 3.8×10^{14} moles carbon fixed per year. Phosphorous is generally assumed to be the limiting nutrient for marine productivity

of geologic timescales (Tyrell 1999) and nitrogen is assumed to be abundant in upwelling waters prior to the evolution of the aerobic nitrogen cycle (Canfield et al. 2010). Dissolved inorganic carbon is assumed to not be limiting for oxygenic photosynthesis, as DIC concentrations are thought to have been as high or higher than modern (Fischer et al. 2009). As a result, DIC would therefore have been greater than 2 mM in seawater, making these concentrations far in excess of other relevant redox-active dissolved species, and thus DIC concentrations do not place a strong limit on oxygenation in so far as coeval redox buffers like ferrous iron are concerned. Burial of organic carbon is assumed to be a similar fraction of primary productivity to the modern, based on Berner (1982), of about 0.32%. Burial rates are considerably higher in modern anoxic basins—e.g. 2% in the Black Sea (Arthur et al. 1994), and 22% in Lake Matano (Crowe et al. 2011), and consequently this approach yields a highly conservative estimate because this parameter exerts strong control on the rate of oxygenation of the oceans and atmosphere.

Model calculations. The model topology consists of a simple box model with initial pools of reduced compounds (Figure 2) that then through numerical simulations are both titrated by O_2 and resupplied. The algorithm is iterated through time, at each step adding a flux of reduced compounds from geological processes as detailed below and in Figure 3 and adding O_2 according to the model parameter GPP. Reduced compounds are consumed by reaction with O_2, and additional oxidizing power is introduced through the escape of hydrogen and photolysis of methane. An example function for tracking a generic reduced compound X is: $X(t)=X(t-1)+X_f*t$, with X(0) equal to the starting pools in Figure 2, X_f is

the flux given in Figure 3, and t is the time step over which the model is iterated. This form of equation is used for reduced compounds with the exceptions of methane and hydrogen, which are taken as functions following the modified form: $X(t)=X(t-1)+X_f*t*e^{-c*t}$, where c is a constant based on estimated lifetimes for these compounds in the Archean atmosphere (Catling et al. 2001). Oxygen is then added to the system according to the chosen GPP value, and allowed to react with each compound. The model iterates through each compound as long as oxygen remains in the system, oxidizing them in a stoichiometric ratio as: $X(t)=X(t)-NetO_2(t)/X_s$, where Xs is the stoichiometry of the reaction and $NetO_2$ is the number of moles of free O_2, beginning each step as $NetO_2(t)=NetO_2(t-1)+GPP*t$ but consumed through each step of reaction with reduced compounds: $NetO_2(t)=NetO_2(t)-X(t)*X_s$. This proceeds at each step until either all O_2 is consumed or all reduced compounds are oxidized, and the algorithm proceeds to the next step, allowing remaining O_2 or reduced compounds to accumulate and persist to the next step. An example output of these calculations is given in Figure 4, in terms of total O_2 equivalents, equivalent to: O_2 Equivalents$=\Sigma X(t)*X_s$.

Results

Under the parameters outlined above, the fluid Earth is oxidized (reduced pools consumed and oxygen begins to accumulate) in about 57 kyr (e.g. Figure 4). This is rapid from the perspective of geological time. While there are uncertainties in these calculations, this suggests that redox buffers were not sufficient to delay oxygenation for hundreds of millions or billions of years. Varying parameters can lead to estimates of oxygenation a few

times faster or slower (scenarios summarized in Table 1), but any realistic combination of values results in oxygenation within at most a few hundreds of thousands of years. Oxygen is then able to accumulate sufficiently to concentrations sufficient to largely eliminate mass independent fraction of atmospheric sulfur-bearing species (e.g. Pavlov and Kasting 2002) within a hundred or so years.

Despite substantial effort modeling the relationships between the core biogeochemical cycles and atmospheric oxygen, little prior work has explicitly addressed the timescales for oxygenation following the evolution of oxygenic photosynthesis. Previous studies that constructed mathematical models of the oxygen cycle typically employed flux estimates that were considered geologically relevant for the problem at hand (e.g. Berner 2006). While this approach is appropriate for the common geological timescales of integration, it smooths over many multiscale processes that have foundationally faster dynamics. For example, it is widely appreciated that burial and weathering (plus volcanic) fluxes are in balance on million year timescales for an atmosphere of ~20% O_2 by volume. (e.g. Lasaga and Ohmoto 2002). However embedded within this, photosynthesis and respiration meet to balance each other on nearly annual timescales. Integrating the oxygen cycle at million year time steps implicitly removes these dynamics. It was our goal with this work here to capture some of this. It is important to note, however, that our results still conform to a number of prior findings regarding the oxygen cycle. Goldblatt and colleagues (2006) showed that the balance between photosynthetic oxygen production and geologic flux of reduced compounds is fundamental to determining steady state oxygen concentrations and that oxygenation of the fluid Earth

would proceed rapidly in less than 150 kyr, a conclusion similar to the model results here. Additionally a recent study that focused on O_2 concentrations through Proterozoic time found that an anoxic or low-oxygen atmosphere is unstable given expectations of an oxygenic photosynthetic biosphere, and that such a system will rapidly converge to a relatively high oxygen atmosphere (Laakso and Schrag 2014). Though arrived at in a different way, the results from that study concur with our findings.

Uncertainties. As discussed above, the most significant variable in this model concerns the fate of fixed organic carbon. As most organic carbon fixed in the modern ocean is quickly respired, the balance of O_2 in the atmosphere can be quickly perturbed if burial rates are altered, and uncertainty in burial efficiency introduces a large degree of uncertainty into our model results. Other factors introduce lesser uncertainty.

Another significant parameter is the proportion of carbon that is remineralized through methanogenesis. Aerobic respiration consumes O_2, but if organic carbon is instead converted to methane it will be oxidized through a combination of methanotrophy—a direct or indirect sink of O_2—as well as photolysis, which is largely O_2-neutral in the Archean atmosphere. As a result, higher methanogenic fluxes actually contribute somewhat to the more rapid accumulation of O_2 in the atmosphere as they provide a sink for reducing equivalents that can be uncoupled from O_2. With a reasonable methanogenic flux of 50% of remineralized carbon [comparable to minimum estimates for ferruginous Lake Matano (Crowe et al. 2011), and those estimated for Archean basins by Hayes (1994)], the oceans and atmosphere achieve oxygenation within ~17 kyr. A more modest methanogenic flux of

10%, well below that of Lake Matano but comparable to that estimated for the pre-oxygenic photosynthesis Archean Earth system (Kharecha et al. 2005) results in oxygenation in about 57 kyr.

While these model calculations draw the best available constraints from theory and analog sites, the Archean Earth is largely unconstrained in many respects, so additional data could change the results. For instance, if early Cyanobacteria were more sensitive to oxygen than anticipated (e.g. by a delayed evolution of oxygen detoxification enzymes), this could in principle limit the local rates of gross primary production, though it remains unclear how significant this would be to the conclusions of our model as we assume a limitation based on nutrient availability rather than growth rate.

Although hydrothermal activity, volcanic outgassing, and crustal production rates are minor components of this model, if their rates are far higher as suggested by some authors (e.g. Kump and Barley 2007) this could delay oxygenation somewhat (though a doubling of either factor affects the time to oxygenation by no more than a few thousand years). However, because of the large difference in magnitude between primary productivity and these geological fluxes they are unlikely to impede oxygenation over geological timescales without unexplained variability through time. Moreover, it is reasonable that Archean tectonic and outgassing rates were more sluggish in the past due to thicker crust (e.g. Korenaga 2006)—even the sign of differences in these rates is uncertain.

Reaching a new steady state. Eventually, oxygen concentrations will stabilize and the Earth will enter a new steady state. Previous models have attempted to resolve the factors

that control O_2 in a steady state (e.g. Lasaga and Ohmoto 2002, Laakso and Schrag 2014,) but the results tend to differ between models and it is difficult to determine which if any are accurate given our current knowledge of the modern O_2 cycle.

The modeling approach taken here differs in that it does not explicitly evaluate how or when a new steady state will be reached, but some predictions can be made. Aerobic respiration serves as one of the most important negative feedbacks on oxygen: higher O_2 and increased oxygenation of the water column and shallow pore fluids promotes more efficient remineralization, limiting the accumulation of oxygen. Before respiration feedbacks take strong hold, the carbon cycle may be out of balance due to organic carbon burial rates higher than those before or since. Similar logic forms the basis for a class of interpretations of the large C isotope excursions observed in carbonate rocks ($\delta^{13}C$ values about 10 permil higher than Archean and Palaeoproterozoic background values) associated with the rise of oxygen (Martin et al. 2013). The exact timing and duration of these isotopic excursions are not well understood (Rasmussen et al. 2013), but they may have lasted for as much as tens to several hundred millions of years (Karhu and Holland 1996). It was suggested that oxygen may have initially accumulated to higher concentrations than supported by this steady state, leading to an initial overshoot followed by collapse before reaching the Proterozoic steady state concentration (Bekker and Holland 2012, Canfield et al. 2013). If correct, substantial amounts of inorganic carbon must be fixed and buried, possibly requiring additional inputs of carbon to the fluid Earth such as the those derived from the oxidation of siderite (Bachan and Kump 2015).

An additional impact of the rise of oxygen was the collapse of the methane greenhouse, which may have helped sustain a clement Paleoproterozoic climate. The timescales estimated for oxygenation (examples in Table 1) are in many cases sufficiently fast to have outpaced the silicate weathering feedback (e.g. Schrag et al. 2002) and triggered global glaciation—possibly marked by the low latitude glacial strata of the Makganyene Formation (Kopp et al. 2005). It is reasonable to suspect that this event included a wide range of impacts on productivity and atmospheric composition during this time.

Conclusions

From the perspective of the calculations presented here, the characteristic timescales of environmental oxygenation [certainly the timescales required to impact sensitive redox proxies like redox-sensitive detrital grains (Johnson et al. 2014) and mass independent S isotope fractionation (Pavlov and Kasting 2002)] following the evolution of oxygenic photosynthesis are geologically rapid—unless rates of GPP are far lower than typically thought. The only way to prevent geologically rapid oxygenation of Earth surface environments once oxygenic photosynthesis has evolved is to reduce rates of primary productivity by orders of magnitude to lower than the geological fluxes of reduced compounds. It has been shown that extant Cyanobacteria experience a defect in growth rate when exposed to high dissolved iron concentrations—an effect termed anaerobic iron toxicity—suggesting that perhaps iron toxicity may have delayed oxygenation following the evolution of oxygenic photosynthesis (Swanner et al. 2015). However, the

observed growth defect was relatively minor (<50%) at the iron concentrations predicted for Archean seawater, and similar and even more severe toxicity has been shown for heterotrophs exposed to high dissolved iron concentrations (e.g. Touati 2000, Bird et al. 2013). As a result, the impact of iron toxicity on Archean net primary productivity may have been neutral or even positive. Additional physiological investigation of Cyanobacteria will be necessary to identify other potential limits to early productivity. However, no compelling argument is currently available for why Cyanobacteria would remain so unproductive over long timescales—and indeed arguments for Cyanobacterial primary production made on the relative abundance of organic carbon in shales (e.g. Lyons et al. 2014), require an explicit balance in terms of O_2 production and a certain degree of ecosystem dominance. Rates of photosynthesis during the rise of oxygen are also not known with any certainty, however there are good reasons to suspect that these rates were elevated above those that characterized the Archean biosphere. The aforementioned interpretations of ^{13}C-enriched isotope ratio data in many Paleoproterozoic carbonates imply a larger relative flux of sedimentary organic matter associated with the rise of oxygen (Karhu and Holland 1996, Rasmussen et al. 2013)—this is in stark contrast to Archean carbonate platforms (Fischer et al. 2009).

A fundamental reason why oxygenic photosynthesis is viewed as such an important metabolic breakthrough is not simply because it harvests more energy than typical anaerobic metabolisms per photon, but also because the electron donor for oxygenic photosynthesis—water—is not limiting in marine environments. Anoxygenic photosynthetic microbes are commonly limited by the supply of electron donors like

molecular hydrogen or ferrous iron (Kharecha et al. 2005, Fischer and Knoll 2009). The removal of this constraint would allow Cyanobacteria to exploit a huge volume of favorable environments on the Earth surface characterized by abundant water and sunlight, with their productivity limited instead by the availability of fixed nitrogen and phosphate.

The evolution of oxygenic photosynthesis fueled a number of adaptive radiations to fill the newly opened niche space of oxygenic photosynthesis and aerobic respiration (and other new metabolic opportunities from evolving biogeochemical cycles of S and N). Adaptive radiations have long been recognized as an important force in the evolution of plant and animal clades (e.g. Gavrilets and Losos 2009), but more recently have also been observed in microbial populations (e.g. Rainey and Travisano 1998). Due to their short generation times, phenotypic adaptation to novel environmental conditions and genomic evolution in bacteria occurs on short timescales even visible in laboratory experiments (Bennett et al. 1992, Barrick et al. 2009). Adaptive radiations involving increased rates of genetic diversification and adaptation commonly occur following evolutionary innovations and opening up of new niche space (Gavrilets and Vose 2005). Consequently it is reasonable to expect that Cyanobacteria rapidly adapted and became specialized within a variety of environments following the evolution of oxygenic photosynthesis. Importantly, this notion is supported by phylogenetic relationships between extant oxygenic Cyanobacteria, which show a rapid radiation following a longer branch (i.e. larger evolutionary distances) separating them from their closest nonphototrophic relatives (Shih et al. 2015). This radiation likely allowed the Cyanobacteria to quickly adapt to an

oxygenic photosynthetic lifestyle and occupy new niches and environments, with a dominant role in photosynthetic ecosystems.

On the basis of the calculations here and the logic presented above, we suggest that oxygenic Cyanobacteria evolved shortly before the rise of oxygen. This finding is further consistent with recent evidence calling into doubt the syngeneity of Archean biomarkers for Cyanobacteria (French et al. 2015), evidence for a transitional manganese-oxidizing phototroph just before the rise of oxygen (Johnson et al. 2013), as well as molecular analyses that reveal the derived position of oxygenic phototrophs within the Cyanobacteria phylum (Di Rienzi et al. 2013, Shih et al. 2013, Soo et al. 2014, Shih et al. 2015, Johnson et al. 2013b). This also intimates that if geochemical data interpreted as whiffs of oxygen from much earlier sedimentary successions are interpreted correctly, they might reflect sources of oxygen derived from other known abiotic [e.g. hydrogen peroxide produced photochemically (Kasting et al. 1985) or via pyrite (Borda et al. 2001) and potentially concentrated on the surface of glacial ice (Liang et al. 2006) or through atmospheric processes (Haqq-Misra et al. 2011)] and biotic [e.g. oxygen produced via catalase reactions (Blankenship and Hartman 1998) or the dismutation of nitric oxide (Ettwig et al. 2010, Ettwig et al. 2012) or chlorite (van Ginkel et al 1996)] sources. Alternatively, these signals could largely be the result of alteration of sedimentary strata by postdepositional fluids. While chemical imaging approaches exist for evaluating the impacts of post-depositional alteration (e.g. Johnson et al. 2013a, Fischer et al. 2014, Stefurak et al. 2015), a critical issue going forward will be constructing analytical frameworks that are able to carefully consider and avoid false positive detections of oxygenic Cyanobacteria in the early

sedimentary record, particularly as geochemical techniques for O_2 detection continue to become more sophisticated and more sensitive.

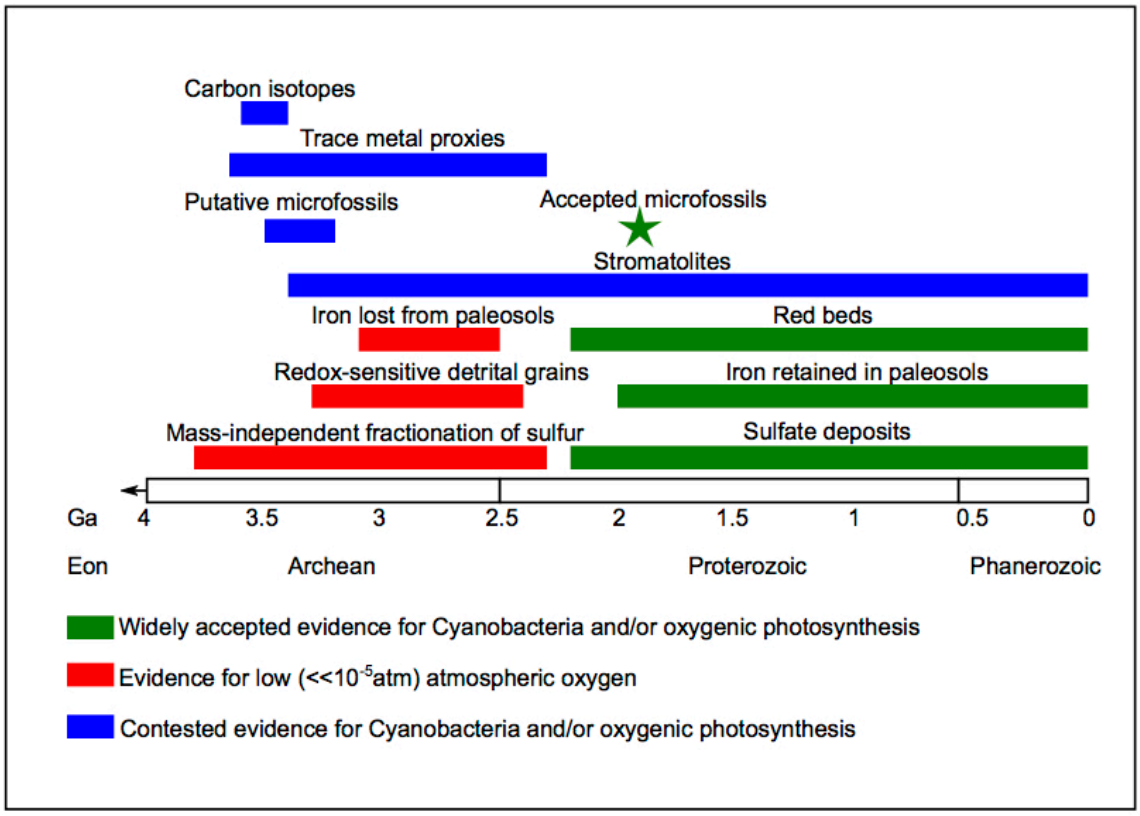


Figure 1: Earth history timeline with key geological and geochemical records of oxygen and Cyanobacteria noted. The rise of oxygen has been dated to *ca.* 2.35 Gya based on the disappearance of mass-independent fractionation of sulfur (Rasmussen et al. 2013). This is consistent with other proxies for a transition in the oxygenation state of the atmosphere at this time such as the disappearance of redox-sensitive detrital grains (Johnson et al. 2014), the behavior of iron in paleosols and red beds (Holland 1984), and the appearance of evaporative sulfate mineral deposits (Chandler 1988). The oldest widely accepted cyanobacterial microfossils are those of *Eoentophysalis sp*. in rocks *ca*. 1.9 Ga. Also noted are several controversial pieces of

evidence for Cyanobacteria and/or oxygen in older rocks. These include putative microfossils at *ca.* 3.5 Ga (Schopf 1993); stromatolites, with a record beginning *ca.* 3.4 Ga (Hofmann et al. 1999, Allwood et al. 2006); carbon isotope values in graphite and kerogen (Schidlowski et al. 1979, Rosing 1999); and various trace metal proxies interpreted as recording local or transient oxygen enrichments between *ca.* 3.7 and 2.5 Ga (Rosing and Frei 2004, Anbar et al. 2007, Planavsky et al. 2014, Crowe et al. 2013).

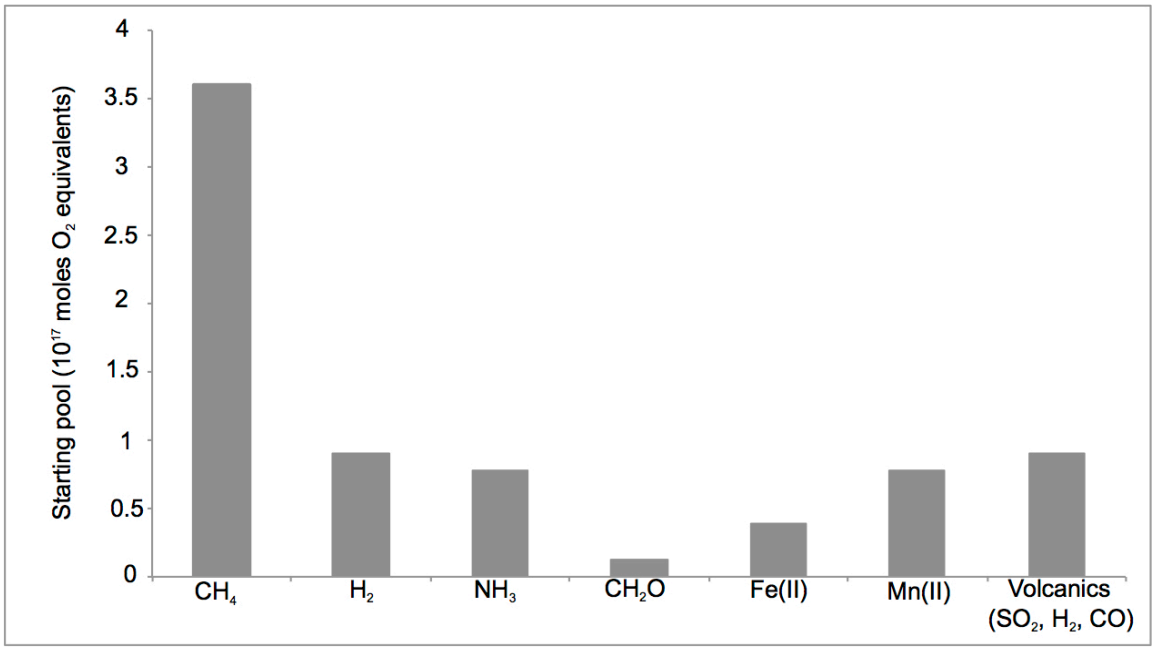


Figure 2: Estimated starting pools of reduced compounds given as moles of O_2 equivalents (e.g. 1 mole of Fe^{2+} is 0.25 moles of O_2 equivalents, as it requires 0.25 moles of O_2 to oxidize 1 mole of Fe^{2+} to Fe^{3+}). H_2 and CH_4 are estimated at 0.001 atm (Pavlov et al. 2000). Fe^{2+} and Mn^{2+} are estimated at up to 120 micromolar (Holland 1984). Ammonia is estimated at 80 micromolar (Canfield et al. 2010). While dissolved iron in the oceans is frequently cited as buffering the Archean Earth against

oxidation, the relatively low concentrations of iron and its oxidation stoichiometry make it relatively insignificant compared to other compounds like methane.

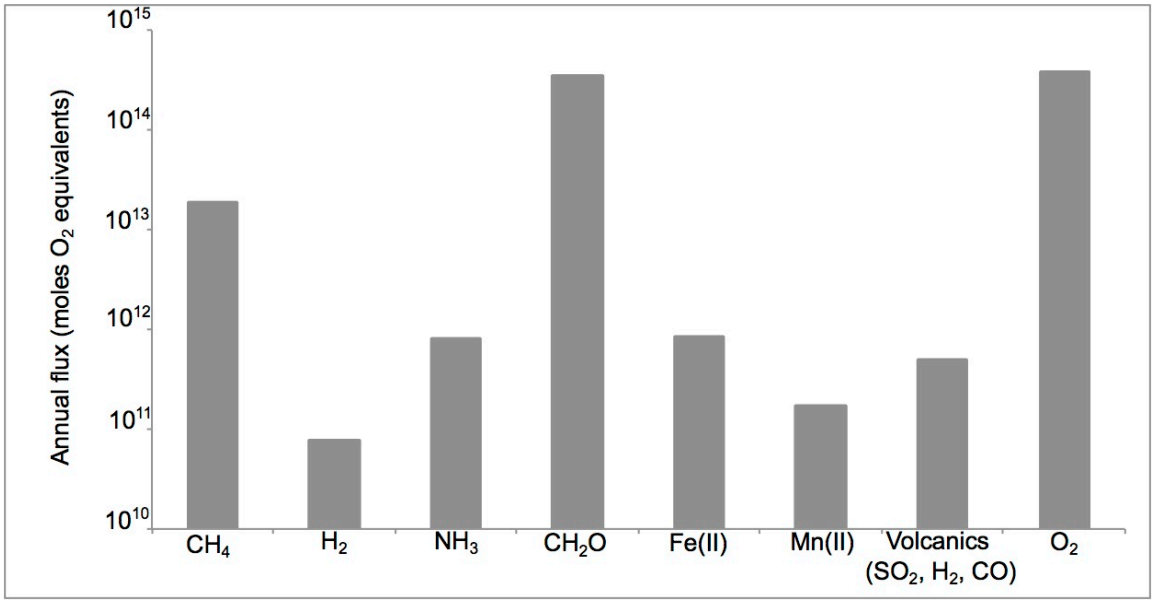


Figure 3: Fluxes of reduced compounds and O_2 given as moles of O_2 equivalents per year. Volcanic gases (Holland 2002), Fe^{2+} (Lécuyer and Ricard 1999), hydrothermal fluxes of other compounds derived from modern fluxes collected by Elderfield and Schultz (1996) and using the high estimate of 10x modern hydrothermal fluxes from Des Marais (2000). O_2, organic carbon, and methane fluxes are variables in our model, and conservative estimates (Scenario A in Table 1) are given here. Primary production fluxes dwarf fluxes of reduced compounds, making them a dominant control on the model results.

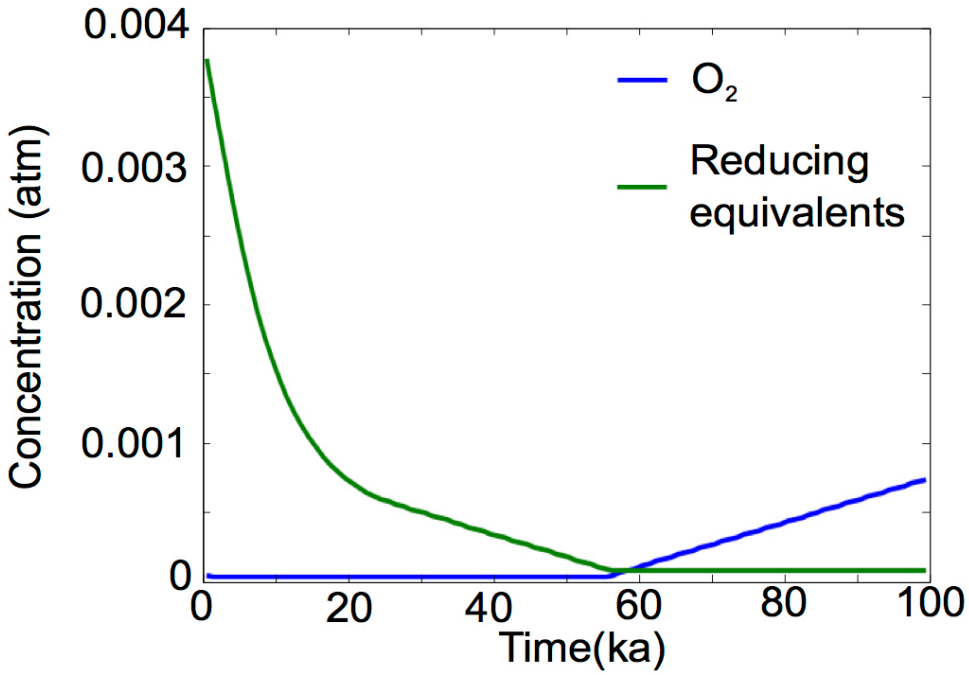


Figure 4: Example calculation, plotting total O_2 equivalents and atmospheric O_2 through time. Data plotted is for Scenario 1 (summarized in Table 1), a conservative estimate of oxygenation time based on the best estimates for Archean NPP (10% of modern marine values based on phosphorous limitation, per Bjerrum and Canfield, 2002), Archean methanogenic rates (0.1 of NPP, per Kharecha et al. 2005), and modern burial rates (0.32% of NPP, per Berner 1982). Once oxygenic photosynthesis is initiated, net O_2 (that remaining after reaction with annual reductant fluxes) begins to consume starting pools of reduced compounds. These are depleted after ~57 kyr, at which time free O_2 accumulates.

Table 1: Model solutions for a range of potential conditions.

Scenario:	**GPP (moles C, O_2/year):**	**Burial fraction:**	**Methanogenic fraction:**	**Oxygenation timescale (ka):**	**Notes:**
A	$3.8x10^{14}$	0.0032	0.1	57	Conservative estimate based on predicted Archean productivity and modern carbon cycle remineralization rates, scenario output shown in Figure 4.
B	$3.8x10^{14}$	0.22	0.5	7	Reasonable Archean estimates based on Lake Matano as an analog
C	$3.8x10^{14}$	0.22	0	8	Scenario B with no methanogenesis
D	$0.19x10^{14}$	0.22	0.5	65	Scenario B with low productivity
E	$3.8x10^{14}$	0.25	0	7	Metabolically-limited (no methanogenesis, inefficient remineralization)

References:

1. Allwood, AC et al. 2006. Nature 441 pp714-718.
2. Anbar, AD and AH Knoll. 2002. Science 297 pp1137-1142.
3. Arthur, MA et al. 1994. Global Biogeochemical Cycles 8 pp195-217.
4. Bachan, A and LR Kump. 2015. PNAS 121 pp6562-6567.
5. Barrick, JE et al. 2009. Nature 461 pp1243-1247.
6. Battistuzzi, FU et al. 2004. BMC Evolutionary Biology 4:44.
7. Bekker, A et al. 2004. Nature 427 pp117-120.
8. Bekker, A and HD Holland. 2012. Earth and Planetary Science Letters 317-318 pp295-304.
9. Bender, M et al. 1994. Global Biogeochemical Cycles 8 pp363-376.
10. Bender, M et al. 1999. Deep-Sea Research I 46 pp637-654.
11. Bennett, AF et al. 1992. Evolution 46 pp16-30.
12. Berner, RA. 1982. American Journal of Science 282 pp451-473.
13. Berner, RA. 1989. Glob. Planet. Change 1 pp97-122.
14. Berner, RA. 2006. Geochim. Cosmochim. Acta 70 pp5653-5664.
15. Bird, LJ et al. 2013. Appl. Environ. Microbiol. 79 pp3619-3627.
16. Bjerrum, CJ and DE Canfield. 2002. Nature 417 pp159-162.
17. Blankenship, RE and H Hartman. 1998. Trends Biochem. Sci. 23 pp94-97.
18. Borda, MJ et al. 2001. Astrobiology 1(3) pp283-288.

19. Bosak, T et al. 2007. Geobiology 5 pp119-126.

20. Brasier, MD et al. 2002. Nature 416 pp76-81.

21. Brock, TD. 1978. Thermophilic Microorganisms and Life at High Temperatures. Springer-Verlag New York.

22. Brocks, JJ et al. 1999. Science 285pp1033-1036.

23. Buick, R. 1992. Science 255 pp74-77.

24. Canfield, DE. 2005. Annual Review of Earth and Planetary Sciences 33pp1-36,

25. Canfield, DE et al. 2010. Science 330 pp192-196.

26. Canfield, DE et al. 2013. PNAS 110(42)pp16736-16741.

27. Catling, DC et al. 2001. Science 293 pp839-843.

28. Chandler, FW. 1988. Carbonates and Evaporites 3 pp75-94.

29. Cloud, P. 1973. Economic Geology 68 pp1135-1143.

30. Crowe, SA et al. 2008. PNAS 105 pp15938-15943.

31. Crowe, SA et al. 2011. Geobiology 9 pp61-78.

32. Crowe, SA et al. 2013. Nature 501 pp535-538.

33. Des Marais, DJ et al. 1992. Nature 359 pp606-609.

34. Des Marais, DJ. 2000. Science 289 pp1703-1705.

35. Di Rienzi, S et al. 2013. eLife 2 pp1-25.

36. Ducluzeau, AL et al. 2014. J. R. Soc. Interface 11.

37. Elderfield, H and A Schultz. 1996. Annu. Rev. Earth Planet. Sci 24 pp191-224.

38. Ettwig, KF et al. 2010. Nature 464 pp543-548.

39. Ettwig, KF et al. 2012. Front. Microbiol. 3 pp1-8.

40. Falkowski, PG. 2011. Photosynthesis Research 107 pp7-10.

41. Farquhar, J et al. 2000. Science 289 pp756-758.

42. Field, CB et al. 1998. Science 281 pp237-240.

43. Fischer, WW and AH Knoll. 2009. Bull. Geol. Soc. Am. 121 pp222-235.

44. Fischer, WW et al. 2009. Precambrian Research 169pp15-27.

45. Fischer, WW et al. 2014. PNAS 111 pp5468-5473.

46. French, KL et al. 2015. PNAS 201419563.

47. Gaillard, F et al. 2011. Nature 479 pp229-232.

48. Gavrilets, S and A Vose. 2005. PNAS 102 pp18040-18045.

49. Gavrilets, S and JB Losos. 2009. Science 323 pp732-737.

50. Glasspool, IJ and AC Scott. 2010. Nature Geoscience 3(9) pp627-630.

51. Gribaldo, S et al. 2009. Trends Biochem Sci. 34 pp375-381.

52. Grotzinger, JP and AH Knoll. 1999. Annu. Rev. Earth Planet. Sci. 27 pp313-358.

53. Guo, Q et al. 2009. Geology 37 pp399-402.

54. Haqq-Misra, J et al. 2011. Astrobiology 11 pp293-302.

55. Hartnett, H et al. 1998. Nature 391 pp572-575.

56. Hayes, JM. 1994. In: Early Life on Earth. Nobel Symposium No. 84 pp.220-236. Columbia U.P., New York.

57. Helz, GR et al. 2011. Chem. Geol. 284 pp323-332.

58. Hofmann, HJ. 1976. Journal of Paleontology 50 pp1040-1073.

59. Hofmann, H. et al. 1999. GSA Bull. 111 pp1256-1262.

60. Hoffman, PF. 2013. *Chem. Geol.* 362pp143-156.

61. Holland, HD. 2002. Geochimica et Cosmochimica Acta 66(21) pp3811-3826.

62. Holland, HD. 1984. The chemical evolution of the atmosphere and oceans. Princeton University Press.

63. Hotinski, RM et al. 2000. Paleoceanography 15 pp267-279.

64. Hotinski, RM et al. 2001. Geology 29 pp7-10.

65. Johnson, JE et al. 2013. PNAS 110 pp11238-11243.

66. Johnson, JE et al. 2013b. PNAS 110 pp E4119–E4120.

67. Johnson, JE et al. 2014. Geological Society of American Bulletin 126 pp813-830.

68. Jones, C et al. 2011. Biogeosciences 8 pp2977-2991.

69. Karhu, JA and HD Holland. Geology 24 pp867-870.

70. Kasting, JF et al. 1985. Journal of Geophysical Research: Atmospheres 90 pp10497-10510.

71. Kharecha, P et al. 2005. Geobiology 3 pp53-76.

72. Kirschvink, JL and RE Kopp. 2008. Phil. Trans. R. Soc. B 363 pp2755-2765.

73. Knoll, AH and S Golubic. 1992. In: Early Organic Evolution pp450-462. Springer Berlin Heidelberg.

74. Kopp, RE et al. 2005. PNAS 102(32) pp11131-11136.

75. Korenaga, J. 2006. Archean Geodynamics and Environments pp7-32.

76. Kump, LR et al. 2001. Geochemistry, Geophysics, Geosystems 2.1.

77. Kump, LR and ME Barley. 2007. Nature 448pp1033-1036.

78. Lasaga, AC and H Ohmoto. 2002. Geochim. Cosmochim. Acta 66 pp361-381.

79. Lécuyer, C and Y Ricard. 1999. Earth and Planetary Science Letters 165 pp197-211.

80. Liang, MC et al. 2006. PNAS 103 pp18896-18899.

81. Lyons, TW et al. 2014. Nature 506pp307–315.

82. Maynard, JB. 2010. Economic Geology 105 pp535-552.

83. Martin, AP et al. 2013. Earth-Science Reviews 127 pp242-261.

84. Morford, JL et al. 2012. Chem. Geol. 324-325 pp73-86.

85. Nägler, TF et al. 2011. Chem. Geol. 289 pp1-11.

86. Owan, T et al. 1979. Nature 277 pp640-642.

87. Papineau, D et al. 2007. Earth Planet. Sci. Lett. 255 pp188-212.

88. Pavlov, AA and JF Kasting. 2002. Astrobiology 2(1) pp27-41.

89. Pavlov, AA et al. 2001. Geology 29(11) pp1003-1006.

90. Pavlov, AA et al. 2000. Journal of Geophysical Research 105 pp11981-11990.

91. Paytan, A et al. 1996. Science 274 pp1355-1357.

92. Planavsky, NJ et al. 2014. Nature Geoscience 7 pp283-286.

93. Rainey, PB and M Travisano. 1998. Nature 394 pp69-72.

94. Rashby, SE et al. 2007. PNAS 104 pp15099–15104.

95. Rasmussen, B et al. 2013. Earth Planet. Sci. Lett. 382 pp173-180.

96. Rasmussen, B et al. 2013b. Geology 41 pp435-438.

97. Ricci, JN et al. 2015. Geobiology 13 pp267-277.

98. Robinson, JM. 1990. Geology 15 pp607-610.

99. Rosing, MT. 1999. Science 283 pp674-676.

100. Rosing, MT and R Frei. 2002. Earth and Planetary Science Letters 217 pp237-244.

101. Sadler, PM. 1981. The Journal of Geology 89 pp569-584.

102. Saito, MA et al. 2003. Inorganica Chimica Acta 356 pp308-318.

103. Schidlowski, M et al. 1975. Precambrian Res. 2 pp1-69.

104. Schidlowski, M et al. 1979. Geochim. Cosmochim. Acta 43 pp189-199.

105. Schidlowski, M 1983. Precambrian Research 20 pp319-335.

106. Schopf, JW. 1993. Science 260 pp640-646.

107. Schrag, DP et al. Geochemistry Geophys. Geosystems 3 pp1-21.

108. Severinghaus, J et al. 2009. Science 324pp1431-1434.

109. Shih, PM et al. 2013. PNAS 110 pp1053-1058.

110. Shih, PM. 2015. Curr. Biol. 25R192–R193

111. Shih, PM et al. 2015. PNAS. Submitted.

112. Soo, RM et al. 2014. Genome Biol. Evol. 6 pp1031-1045.

113. Stolper, DA et al. 2010. PNAS 107(44) pp18755–18760.

114. Suwa, M and ML Bender. 2008. Geophys. Res. Atmos. 113 pp1-9.

115. Stefurak, EJT et al. 2014. Geochim. Cosmochim. Acta 150 pp26-52.

116. Tice, MM and DR Lowe. 2004. Nature 431 pp549-552.

117. Touati, D. 2000. Arch. Biochem. Biophys. 373 pp1-6.

118. Tyrrell, T. 1999. Nature 400(6744) pp525-531.

119. Van Dinkel, CG et al. 1996. Archives of Microbiology 166 pp321-326.

120. Walker, JCG et al. 1983. In: Earth's earliest biosphere: Its origin and evolution pp260-290. Princeton, NJ, Princeton University Press.

121. Welander, PV et al. 2010. PNAS 107 pp8537-8542.

122. Wellman, CH et al. 2003. Nature 425 pp282-285.

123. Williams, HM et al. 2012. Earth and Planetary Science Letters 321-322 pp54-63.

Chapter 3

A COMPLEX HISTORY OF PHOTOTROPHY REVEALED BY NOVEL CHLOROFLEXI LINEAGES

Ward, Lewis M., James Hemp, Patrick M. Shih, Shawn E. McGlynn, and Woodward W. Fischer. In preparation.

Abstract:

The antiquity of photosynthesis is a point of substantial contention, with competing hypotheses for either its presence in the last universal common ancestor and gene loss in most lineages, or a much later origin followed by extensive horizontal gene transfer into the extant phototrophic clades. Selecting between these alternatives requires constraining whether photosynthesis-associated genes have been inherited vertically from a common photosynthetic ancestor, or whether they have been horizontally transferred. While the phylogenetic relationships between the bacterial phyla are unclear, making the order of acquisition of photosynthesis uncertain, much greater resolution is retained within phyla, where more recent cases of vertical inheritance or horizontal gene transfer can be tested. Here, we report several new draft genome sequences from within the Chloroflexi phylum that form a metabolically diverse, monophyletic clade sister to the Anaerolineae class that we term *Candidatus* Thermofonseaa. This class includes two independently phototrophic lineages. Comparison of organismal (based on conserved ribosomal proteins) and phototrophic gene (based on reaction center and bacteriochlorophyll synthase) trees demonstrate that these lineages acquired phototrophy via horizontal gene transfer from the phototrophic Chloroflexia class. The extent of horizontal gene transfer of phototrophy, as

well as other metabolic pathways, within the Chloroflexi is indicative of a broader role for horizontal transfer in the distribution of metabolic traits throughout the bacteria, challenging assumptions about metabolism in the last universal common ancestor and subsequent innovation and diversification of metabolic pathways.

Introduction:

Anoxygenic photosynthesis is among the oldest and most important metabolic inventions in the history of life on Earth. Anoxygenic photosynthesis was a critical driver of primary productivity on the early Earth, and it gave rise to oxygenic photosynthesis, which revolutionized biology and geochemistry and ultimately fueled the rise of complex life. Multiple hypotheses exist for the origin and subsequent evolution of anoxygenic photosynthesis, but little is known with certainty. To date, chlorin-based phototrophy has been identified in seven bacterial phyla: the Cyanobacteria, Chlorobi, Chloroflexi, Acidobacteria, Heliobacteria, Gemmatimonadetes, and Proteobacteria (Figure 1). Of these, only one—the Cyanobacteria—possesses two photosystems coupled to perform oxygenic photosynthesis. The others perform anoxygenic phototrophy and possess only a single reaction center, either Type 1 (Chlorobi, Heliobacteria, and Acidobacteria) or Type 2 (Proteobacteria, Gemmatimonadetes, and Chloroflexi). While it has been suggested that photosynthesis was present in the last common ancestor of the bacteria (Woese et al. 1985, Woese 1987), followed by extensive loss in most lineages, this idea remains controversial. An alternative scenario involves a later origin of photosynthesis, followed by multiple instances of horizontal gene transfer (HGT) resulting in the modern distribution of photosynthesis (e.g. Raymond et al. 2002).

A test of these alternatives is to compare the organismal phylogenies of phototrophic bacteria to the gene phylogenies of photosynthesis genes—concordance of the trees would indicate shared ancestry, while discrepancies between them would indicate a history of horizontal gene transfer (Doolittle 1986). While the structure of the bacterial tree of life is debated (e.g. Woese 1987, Williams et al. 2013, McInerney et al. 2014), intra-phylum organismal relationships are generally robust despite uncertainty in relationships between phyla (Pace 2009). As a result, the history of photosynthesis within a phylum is more straightforward to assess than it is for the bacteria as a whole. If a major role for horizontal gene transfer can be demonstrated within a particular phylum, the HGT-driven phototrophy hypothesis will be strengthened, whereas a concordance of organismal and gene trees would be more consistent with an ancient origin and vertical inheritance of the metabolism. While tests of this kind have been made previously in the Proteobacteria, suggesting intra-phylum horizontal gene transfer (Igarashi et al. 2001, Nagashima & Nagashima 2013), this has not previously been possible in other phototrophic phyla due to the limited diversity of phototrophic members within each. However, new opportunities for assessing the history of photosynthesis within the Chloroflexi phylum has recently been made possible by the discovery of novel lineages of phototrophs within this phylum.

The Chloroflexi (i.e. Green Nonsulfur Bacteria) are a phylum of primarily gliding, filamentous bacteria possessing a wide diversity of metabolisms and ecological roles, but are best known as photoheterotrophs (Overmann 2008). Chloroflexi have been shown to be diverse and abundant in a range of environments (e.g. marine sediments and groundwater, Inagaki et al. 2003, Hug et al. 2013). Despite their environmental richness revealed by

culture-independent surveys, most described Chloroflexi belong to a few subclades isolated from hot springs (Yamada and Sekiguchi 2009), including the anoxygenic phototrophic *Chloroflexus* (Pierson and Castenholz 1974, Hanada et al. 1995). Phylogenetic analysis of the phototrophy genes of Chloroflexi has suggested that anoxygenic photosynthesis in this group predates the evolution of oxygenic photosynthesis in Cyanobacteria (Xiong et al. 2000), implying that this group is ancient and therefore a good candidate for investigating questions of early Earth history. Recent culture- and sequence- based efforts have expanded the known taxonomic and metabolic diversity of the Chloroflexi phylum (e.g the Ardenticatenia class, capable of nitrate- and iron oxide- reduction, Hemp et al. 2015b, Kawaichi et al. 2015). Newly discovered Chloroflexi vary tremendously in their morphology, metabolism, and other traits (Table 1), but are recovered as a monophyletic clade in phylogenetic trees (Figure 2) and have sufficient sequence similarity to be classified as a single phylum (Hanada 2014).

Here, we report draft genomes of three lineages of phototrophic Chloroflexi, including two outside of the classically phototrophic Chloroflexia class, and demonstrate the incongruence of organismal and phototrophic gene trees, suggesting a history of horizontal gene transfer of photosynthesis within this phylum. Moreover, multiple genome bins were recovered for nonphototrophic relatives of two of these novel phototrophic Chloroflexi, forming a new clade sister to the Anaerolineae class of Chloroflexi. This new class-level clade was recovered as several genome bins from Japanese hot spring metagenomes, including members with genes for aerobic and anaerobic respiration as well as both chlorin- and rhodopsin-based phototrophy. The discovery of diverse members of an

entire new class of Chloroflexi from just two Japanese hot springs suggests a large amount of unknown diversity in this phylum that may be recovered by further culture- and sequencing-based efforts.

Methods:

Metagenomic sample collection:

Four metagenomic datasets were recovered from two hot springs in Japan, Jinata Onsen and Nakabusa Onsen. Genome bins labeled JP1 or JP3 were derived from Jinata Onsen, while CP1 and CP2 were derived from Nakabusa Onsen.

Jinata genome bins were assembled from two metagenomes from Jinata Onsen, on Shikinejima Island, Tokyo Prefecture. The geochemistry and microbial diversity of this spring is described in detail elsewhere (Ward et al. 2017a). Jinata Onsen is located at 34.326111N, 139.21E on the island of Shikinejima, Tokyo Prefecture, Japan. Shikinejima is part of the Izu Islands, a chain of volcanic islands that formed in the past 2-3 million years along the northern edge of the Izu-Bonin-Mariana Arc (Kaneoka et al. 1970). The source water of Jinata Onsen emerges anoxic, iron-rich, and gently bubbling from the spring source. Temperatures at the source are ~62°C. This spring water flows into a series of pools that mix progressively more with seawater during high tide, creating a range of geochemical conditions over short spatial and temporal scales as hot, iron-rich, oxygen-poor spring water mixes with cold, sulfate- and oxygen-rich seawater. The metagenomes from which JP1 bins were recovered was sequenced from an iron-oxide rich pool near the spring source (Pool 1), while JP3 genomes were recovered from a Cyanobacteria-rich

microbial mat in Pool 3, the most downstream section of the hot spring before it flows into the open ocean. Dissolved oxygen (DO), pH, and temperature measurements were performed *in situ* using an Exetech DO700 8-in-1 Portable Dissolved Oxygen Meter. Iron concentrations were measured using the ferrozine assay (Stookey 1970) following acidification with 40 mM sulfamic acid to inhibit iron oxidation by O_2 or oxidized nitrogen species (Klueglein and Kappler 2013). At the time of sampling, Pool 1 59°C, pH 5.8, and contained 1.8 mg/L dissolved oxygen and 265 μM Fe^{2+}; Pool 3 at Jinata Onsen was 46°C, pH 6.7, and contained 5.6 m/L dissolved oxygen and 100 μM Fe^{2+}.

Nakabusa genome bins were assembled from two metagenome from hot spring microbial mats from Nakabusa Onsen, located at 36.392429N, 137.748038E in the Japanese Alps near Azumino, Nagano Prefecture. Geochemical and microbiological characterization of Nakabusa Onsen is described in detail elsewhere (Kubo et al. 2011, Ward et al. 2017b). Nakabusa Onsen is a sulfidic, moderately alkaline hot spring with source waters near 70°C. The samples from which the metagenomes were sequenced were of cone-forming microbial mats at two points along the outflow from the hot spring source; Cone Pool 1 (the source of CP1 genomes) was a Chloroflexi-dominated mat near the hot spring source, which at the time of sampling was 48°C and pH 8.1, while Cone Pool 2 (the source of the CP2 genomes) was collected from a cone-forming Cyanobacteria-rich microbial mat several meters downstream, which at the time of sampling was 32°C and pH 8.3.

Samples of microbial mats were collected using sterile forceps and spatulas (~0.25 cm^3 of material). Cells were lysed and DNA preserved in the field using Zymo Terralyzer

BashingBead Matrix and Xpedition Lysis Buffer. Cells were disrupted immediately by attaching tubes to the blade of a cordless reciprocating saw and operating for 1 minute.

Metagenomic sequencing and analysis:

Following return to the lab, DNA was extracted and purified with a Zymo Soil/Fecal DNA extraction kit (Zymo Research, Irvine, CA). DNA was quantified with a Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA) according to manufacturer's instructions following DNA extraction. Purified DNA was submitted to SeqMatic LLC (Fremont, CA) for library preparation and sequencing via Illumina HiSeq technology. Raw sequences were assembeled with MegaHit v. 1.02 (Li et al. 2016) and genome bins constructed using MetaWatt version 3.5.2 (Strous et al. 2012). Genomes were manually screened for genes of interest and uploaded to RAST (Aziz et al. 2008) for overall characterization. Genome bins were assessed for completeness and contamination using CheckM (Parks et al. 2014). Genes of interest (e.g. coding for ribosomal, photosynthesis, and electron transport proteins) were screened against outlier (e.g. likely contaminant) contigs as determined by CheckM using tetranucleotide, GC, and coding density content.

Phylogenetics

Sequences of ribosomal and phototrophy proteins were identified locally with BLAST+ (Camacho et al. 2008), aligned with MUSCLE (Edgar 2004), and alignments manually curated in Jalview (Waterhouse et al. 2009). Phylogenetic trees were calculated using RAxML (Stamatakis 2014) on the Cipres science gateway (Miller et al. 2010). Trees were visualized with Seaview (Gouy et al. 2010) and the Interactive Tree of Life (Letunic and Bork 2016).

Results and discussion:

Our sequencing efforts, including both hot spring metagenomes and sequencing of cultured isolates, have resulted in draft genomes of three new reaction center-containing phototrophic Chloroflexi lineages and eight genome bins which do not encode reaction centers but are associated with a new class-level clade sister to the Anaerolineae (Table 1, Figure 2). *Kouleothrix aurantiaca* represents a so-far monospecific genus within the class Chloroflexia, while JP3_7, CP2_42A, and the other genome bins reported here form a new clade sister to the Anaerolineae. Genome statistics for these bins are reported in Table 1, with summaries of the metabolic proteins encoded by these genomes in Table 2.

Organismal phylogenies of the Chloroflexi phylum, including the novel phototrophs and other genome bins described here, were constructed using the RpoB protein sequence (Figure 2). This protein is a core information processing protein, and is always found as a single copy that has been vertically inherited (Hansmann and Martin 2000), so the RpoB phylogeny should reflect the organismal phylogeny. For genomes in which a 16S gene was recovered, the 16S phylogeny matched the topology of the RpoB tree. In these phylogenies, *Kouleothrix aurantiaca* branched within the Chloroflexia class, basally to the *Roseiflexus* after their divergence from *Chloroflexus*. CP2_42A and JP3_7 were recovered as separate lineages, forming a clade sister to the Anaerolineae along with the other genome bins reported here.

Reaction center protein trees (Figure 3) show *Kouleothrix* in the same position relative to other Chloroflexia as in organismal trees, basal to the *Roseiflexus*, but place

CP2_42A and JP3_7 very differently—with CP2_42A as branching between *Kouleothrix* and *Roseiflexus*, and JP3_7 branching sister to the *Roseiflexus+CP2_42A+Kouleothrix* clade.

Candidatus Thermofonseae, a metabolically diverse novel class of Chloroflexi sister to Anaerolineae

The metagenome bins reported here, together with the "Anaerolineae-like" phototroph recovered from a Yellowstone National Park metagenome (Klatt et al. 2011), form a monophyletic clade sister to the Anaerolineae class in phylogenetic trees based on conserved organismal marker proteins such as RpoB (Figure 2). The members of this class appear to encode diverse heterotrophic metabolic traits, including photoheterotrophy and diverse pathways for both aerobic and anaerobic metabolism.

We propose for this clade the name *Candidatus* Thermofonseae, from the Latin for hot spring, with official classification pending isolation and characterization of at least one member. Genome bins falling within this clade were recovered from all four of our metagenomic datasets from both hot springs. The Thermofonseae genome bins reported here were up to ~96% complete as determined by single copy marker genes, and recovered diverse metabolic capabilities as described above, distinguishing them from their closest relatives, the metabolically-limited Anaerolineae. Of the ten genomes reported here, two include phototrophic reaction centers, two include rhodopsins, two possess partial denitrification pathways, and six contain genes for aerobic respiration (Supplemental Table 2). The Anaerolineae, in contrast, are typically described as obligate anaerobes (e.g. Yamada and Sekiguchi 2009), though genes for aerobic respiration have been recovered in

multiple Anaerolineae genomes (e.g. Pace et al. 2015, Ward et al. 2015a, Hemp et al. 2015c). Phylogenetic analysis of electron transport and respiration genes in the Thermofonseae and Anaerolineae reveal that metabolic protein trees are incongruent with organismal trees, suggesting independent acquisitions of respiration in these two clades (Supplemental Figures). The Thermofonseae utilize a bc complex for respiration, while the Anaerolineae commonly use an Alternative Complex III. Furthermore, the Heme Copper Oxidoreductases (HCOs) in these organisms are not closely related. Thermofonseae uses an A family HCO closely related to those of Cyanobacteria, while those in Anaerolineae are closely related to Caldilineae. This suggests that stem lineages of these classes diverged prior to the acquisition of aerobic respiration, followed by diversification after receiving this metabolism through horizontal gene transfer, or alternatively loss and replacement in one lineage.

At least three members of the Thermofonseae (CP2_42A, JP3_13, and CP2_2F) contain rhodopsin genes. Those of JP3_13 and CP2_2F are related to the "Actinorhodopsins" found in *Roseiflexus sp*. RS-1 which is thought to be functional as a light-driven proton pump (Sharma et al. 2008). The rhodopsins found in JP3_13 and CP2_2F have highly similar sequences; given the relatedness of these strains and their rhodopsins, these genes were likely inherited from the last common ancestor of these strains following acquisition via horizontal gene transfer. The rhodopsin encoded in the CP2_42A genome is most closely related to xanthorhodopsin, a proton-pumping rhodopsin shown to use light-harvesting antenna carotenoids (Balashov et al. 2005). Despite the presence of rhodopsins in diverse members of the Chloroflexi including *Roseiflexus*,

Ktedonobacter racemifer, and *Bellilinea caldifistulae* (members of the Chloroflexia, Ktedonobacteraceae, and Anaerolineae classes of the Chloroflexi, respectively), the rhodopsins in each of these Chloroflexi lineages are not closely related, and likely reflect independent acquisitions via horizontal gene transfer from other phyla and not a shared history of rhodopsins in the Chloroflexi phylum. Genes for the copper-containing nitrite reductase NirK are present in bins JP1_20 and CP2_20G, suggesting the potential for at least partial denitrification in these strains.

Anoxygenic phototrophs

Two members of *Ca*. Thermofonseae reported here, CP2_42A and JP3_7, contained genes indicate of anoxygenic phototrophy. The CP2_42A genome bin branches basally to the Anaerolineae class based on organismal trees built based on RpoB and concatenated ribosomal proteins. CP2_42A encodes genes for anoxygenic phototrophy; a type II reaction center (including a fused pufLM and a pufC), bacteriochlorophyll biosynthesis, a cytochrome bc complex, and Alternative Complex III.

JP3_7 is most closely related to the "Anaerolineae-like" phototrophic Chloroflexi assembled from a metagenome from Yellowstone National Park (Klatt et al. 2011), though it is genetically distinct to the species and possibly the genus level (~78% average nucleotide identity). JP3_7 encodes genes for anoxygenic phototrophy; a type II reaction center (including unfused pufL and pufM, and a pufC), bacteriochlorophyll synthesis, a cytochrome bc complex, but no Alternative Complex III. JP3_7 contains unfused PufL and PufM genes, similar to that in *Chloroflexus*.

Although CP2_42A and JP3_7 contain most genes involved in bacteriochlorophyll synthesis, including bchX,Y,Z,P, F,G,I,D, and a bchH homolog, both are missing bchL,N,B,E, or M. Neither CP2_42A nor JP3_7 bin recovered genes for the Calvin cycle or the 3-hydroxyprionate bicycle and so these strains may grow primarily as photoheterotrophs.

While the draft genomes reported here are largely too fragmented to recover informational genes on the same contigs as phototrophy related genes, the rpoB and bchP genes of JP3_7 are collocated on contig 8001, providing strong support for the inference of phototrophy in this lineage from the genome bins produced here.

Horizontal gene transfer of phototrophy within the Chloroflexi:

The position of *Kouleothrix* in both organismal and gene trees is consistent with a vertical inheritance of phototrophy from the last common ancestor of the *Roseiflexus+Chloroflexus* clade after its divergence from the nonphototrophic *Herpetosiphon*. However, the other two phototrophic Chloroflexi reported here reveal a more complicated history. These two strains (CP2_42A and JP3_7) branch within *Ca*. Thermofonseae in organismal trees based on conserved vertically inherited proteins like RpoB, quite distant from the other phototrophic Chloroflexi (Figure 2). However, photosynthesis-related gene trees place these strains within the other phototrophic Chloroflexi. CP2_42A branches within the Chloroflexia, basal to the clade of *Roseiflexus* and *Kouleothrix*. JP3_7, however, branches more deeply, sister to the *Roseiflexus+Kouleothrix+CP2_42A* clade (Figure 3). Furthermore, *Kouleothrix* and CP2_42A have fused pufL and pufM genes, a feature which appears in reaction centers of

Roseiflexus (Youvan et al. 1984, Yamada et al. 2005) and so appears to be a synapomorphy of this lineage of phototrophs, supporting their inclusion at this point in the phototrophy tree to the exclusion of JP3_7 which has unfused pufL and pufM genes.

The incongruence between organismal and gene trees for the novel phototrophic Chloroflexi described here suggests that photosynthesis genes were not vertically inherited from the last common ancestor of the phototrophic Chloroflexi. Instead, the differing branching order of JP3_7 and CP2_42A between RpoB and PufLM trees, along with the presence of a conserved gene fusion within the *Roseiflexus+Kouleothrix+CP2_42A* clade, strongly suggests that horizontal gene transfer has played a role in the current distribution of phototrophy in the Chloroflexi phylum.

In light of these data, the most parsimonious scenario for the evolution of phototrophy within the Chloroflexi requires at least two instances of horizontal gene transfer to have occurred (Figure 4). This scenario involves the acquisition of an unfused Type 2 reaction center in an ancestor of the Chloroflexia after the divergence of *Herpetosiphon*, horizontal gene transfer of this unfused ancestral form from the branch leading to *Roseiflexus* into the JP3_7 lineage, followed by a single pufL+pufM fusion event in an ancestor of *Roseiflexus* and *Kouleothrix*, and a second horizontal gene transfer event of the now fused protein into an ancestor of CP2_42A from the *Roseiflexus* lineage after its divergence from *Kouleothrix*.

Presence of other photosynthesis-related genes

The genome bins for CP2_42A and JP3_7 recover most, but not all, of the bacteriochlorophyll synthesis pathway expected for phototrophic Chloroflexi. These

genomes contain bchX,Y,Z,P,F,G,I,D, and a bchH-like gene, but not bchL,N,B, M, or E. While this may be a result of the incomplete nature of these genomes, the same bacteriochlorophyll synthesis gene complement has been described in the "Anaerolineae-like" phototroph genome bin recovered from a Yellowstone National Park metagenome (Klatt et al. 2011). Some or all of these genes may in fact actually be absent from these genomes, possibly functionally replaced by promiscuous homologs (e.g. bchLNB are homologous to bchXYZ, and chimeras of other homologs of these genes have been demonstrated to be functionally exchangeable, e.g. Wätzlich et al. 2009, Cheng et al. 2005). Isolation and biochemical characterization of the bacteriochlorophyll synthesis pathway in these organisms will be necessary to resolve this possibility. Estimates of the probability of missing the same set of genes from multiple genomes of relatively high (>50%) completeness are incredibly low (<<1%), strengthening interpretations of the absence of these genes from *Ca*. Thermofonseae genomes (Supplemental Information).

It is interesting to note that the vast majority of sequenced phototrophic Chloroflexi utilize Alternative Complex III (Yanyushin et al. 2005) for energy conservation during phototrophy, even to the extent of CP2_42A appearing to have acquired ACIII along with other phototrophy genes (Supplemental Information). However, ACIII was not recovered in the draft genomes for *K. aurantiaca* or JP3_7. This suggests that the use of ACIII for phototrophy may not by a synapomorphy of the phototrophic Chloroflexi, though this will require closure of these genomes and confirmation that ACIII is truly absent and not simply missing from the draft genome. The presence of auracyanin, the electron acceptor of ACIII (Majumder et al. 2013), in JP3_7 is consistent with the ancestral presence of ACIII in this

lineage and either recent loss or failure to recover the gene in the genome bin. Meanwhile, all of the aerobic members of the Thermofonseae encode a bc complex, consistent with other aerobic, nonphototrophic Chloroflexi clades such as Caldilineae and Ardenticatenia.

The history of carbon fixation in the Chloroflexi is complicated, and is not only not congruent with the organismal phylogeny but is not congruent with phototrophy gene phylogenies, suggesting that the light and dark reaction pathways for photosynthesis in the Chloroflexi have undergone independent histories of horizontal gene transfer. While the phototrophic Chloroflexi are well known to possess the 3-hydroxypropionate bicycle for carbon fixation (e.g. Berg 2011), this pathway is absent in the genomes reported here, as well as *Oscillochloris* and *Chlorothrix*. Instead, *Kouleothrix, Oscillochloris,* and *Chlorothrix* possess the Calvin Cycle, while CP2_42A and JP3_7 do not encode and carbon fixation pathways. Comparison of the phylogenies of these organisms and their phototrophy genes does not reveal a clear, consistent history for carbon fixation, with scenarios involving the first phototrophic Chloroflexi possessing the 3-HP bicycle, the Calvin cycle, or no carbon fixation at all being similarly parsimonious.

Conclusions:

The increased diversity of Chloroflexi phototrophs and history of HGT described here are consistent with other metabolic characters in this phylum. Our group has previously sequenced diverse representatives of the Chloroflexi, filling in gaps in the tree (Ward et al. 2015ab, Hemp et al. 2015abc, Pace et al. 2015) in order to better characterize the diversity and distribution of high potential metabolism within this phylum. Our previous efforts to expand the sequenced diversity of the Chloroflexi have revealed a high

degree of previously unrecognized metabolic diversity in this phylum, including high-potential metabolic pathways for aerobic and anaerobic respiration (Ward et al.2015ab, Hemp et al. 2015abc, Pace et al. 2015). The distribution and phylogenies of genes associated with these pathways is diverse and consistent with a history of horizontal gene transfer. Together, these data are consistent with a high degree of metabolic diversity and abundant horizontal gene transfer within the Chloroflexi phylum.

A history of horizontal gene transfer of phototrophy within the Chloroflexi is consistent with that of the Proteobacteria, which records extensive intra-phylum HGT (Igarashi et al. 2001, Nagashima & Nagashima 2013). A single clear case of inter-phylum HGT is also recorded in the presence of a Proteobacteria-derived RC2 in a member of the Gemmatimonadetes (Zeng et al. 2014).

An important question is the relative timing of acquisition of phototrophy in different lineages. If anoxygenic photosynthesis is an ancient metabolism, predating the introduction of molecular oxygen into biology as a result of oxygenic photosynthesis (e.g. Xiong et al. 2000), then at least some lineages should have acquired anoxygenic photosynthesis before the Great Oxygenation Event ~2.3 Gya, and therefore before the ability to respire oxygen. However, for many phototrophic groups, it appears that the acquisition of phototrophy postdated the acquisition of aerobic respiration (Fischer et al. 2016). As a result, the taxonomic affinity of the oldest anoxygenic phototrophs remains unclear. It is possible that photosynthesis originated in a member of a characterized phototrophic clade (e.g. the suggestion that phototrophy may have originated in the Chloroflexi, Oyaizu et al. 1987, or the Cyanobacteria, Mulkidjainian et al. 2006), within a

so-far undiscovered but still extant group, or may in fact have gone extinct. This can best be resolved by continued discovery of new phototrophic groups—an increasingly frequent phenomenon as environmental sequencing efforts continue and improve.

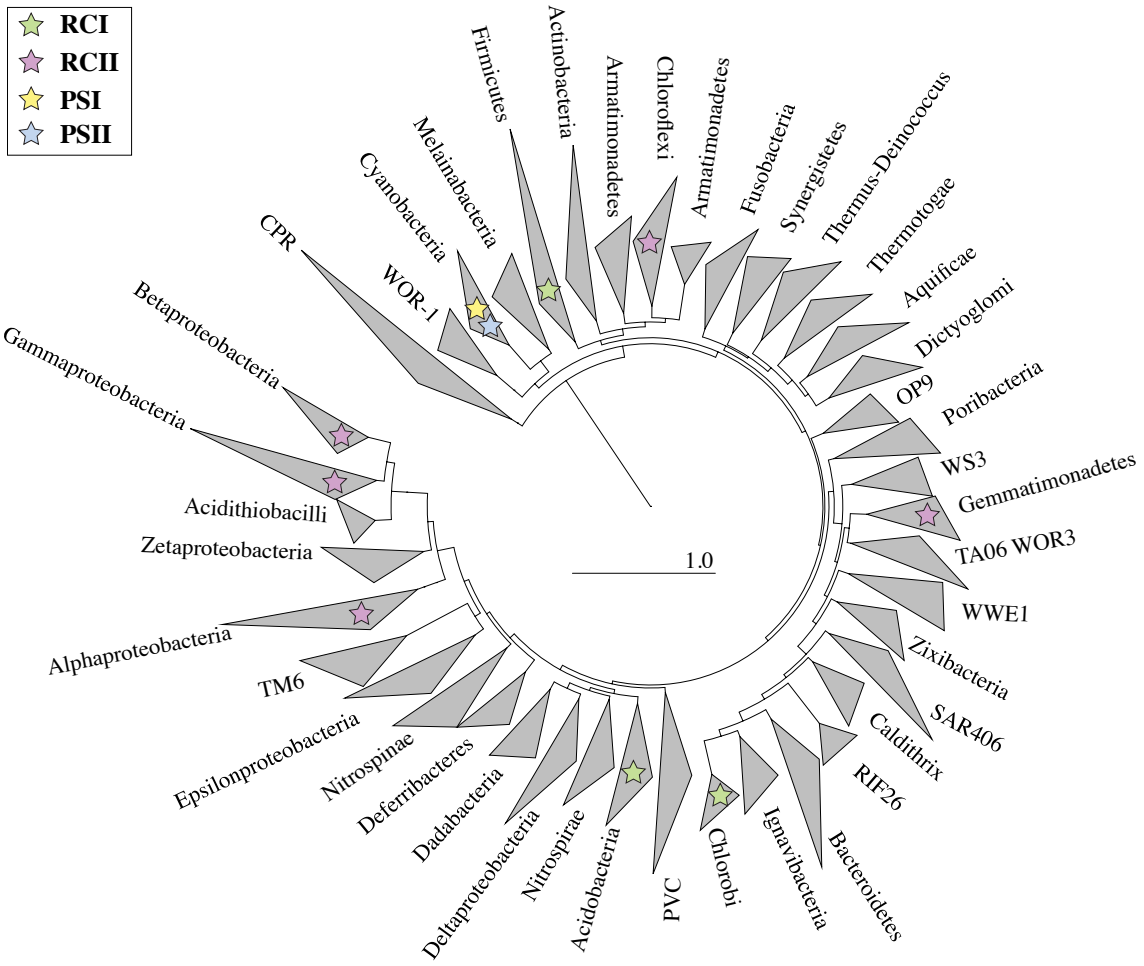


Figure 1: Phylogeny of bacteria based on concatenated ribosomal proteins, with clades that include phototrophs noted with symbols representing their reaction center. Modified from Hug et al. 2016 and visualized using the interactive tree of life (Letunic and Bork 2016).

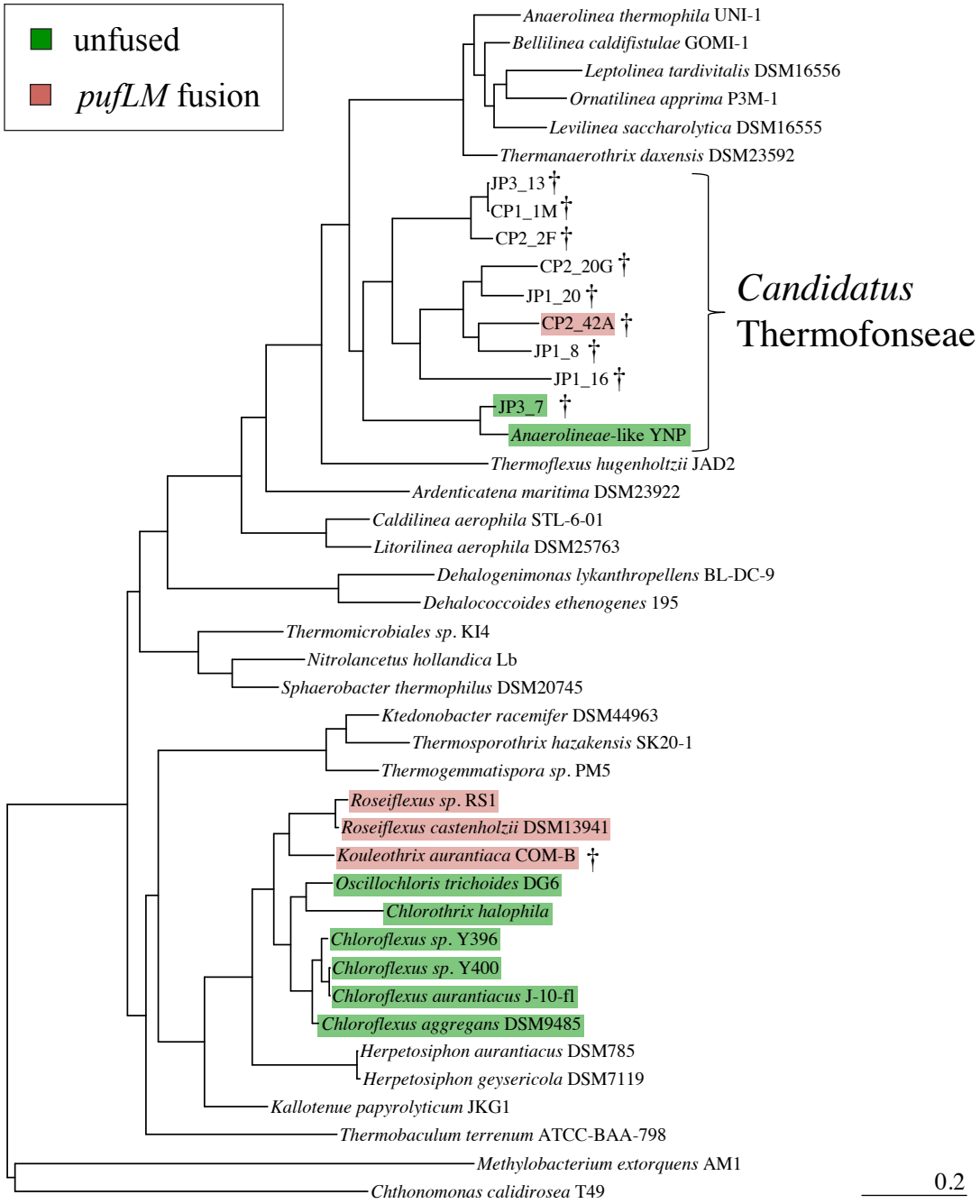


Figure 2: Phylogeny of Chloroflexi based on the RpoB protein sequence, with our newly sequenced strains indicated with daggers, phototrophic strains highlighted (red for fused pufLM, green for unfused), and *Candidatus* Thermofonseae noted.

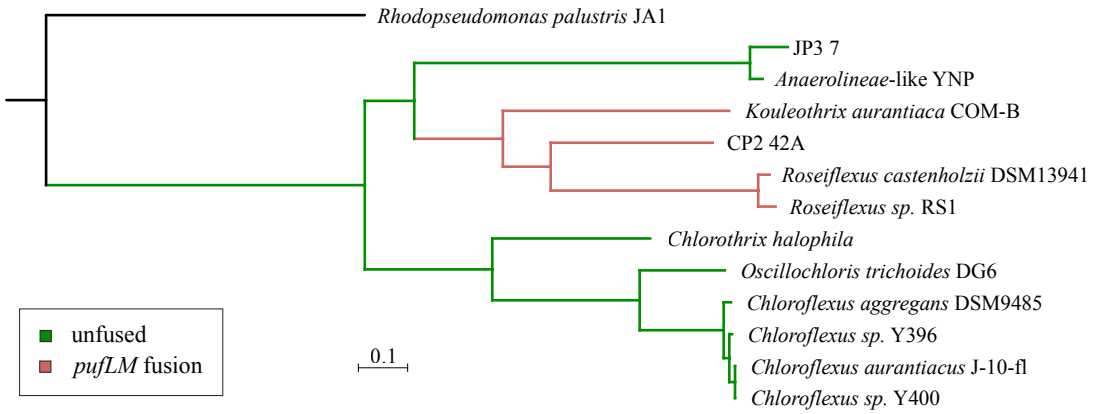


Figure 3: Phylogeny of Type 2 phototrophic reaction center proteins (concatenated PufL and PufM); lineages with fused PufLM proteins are highlighted in red, while lineages with unfused reaction centers are in green.

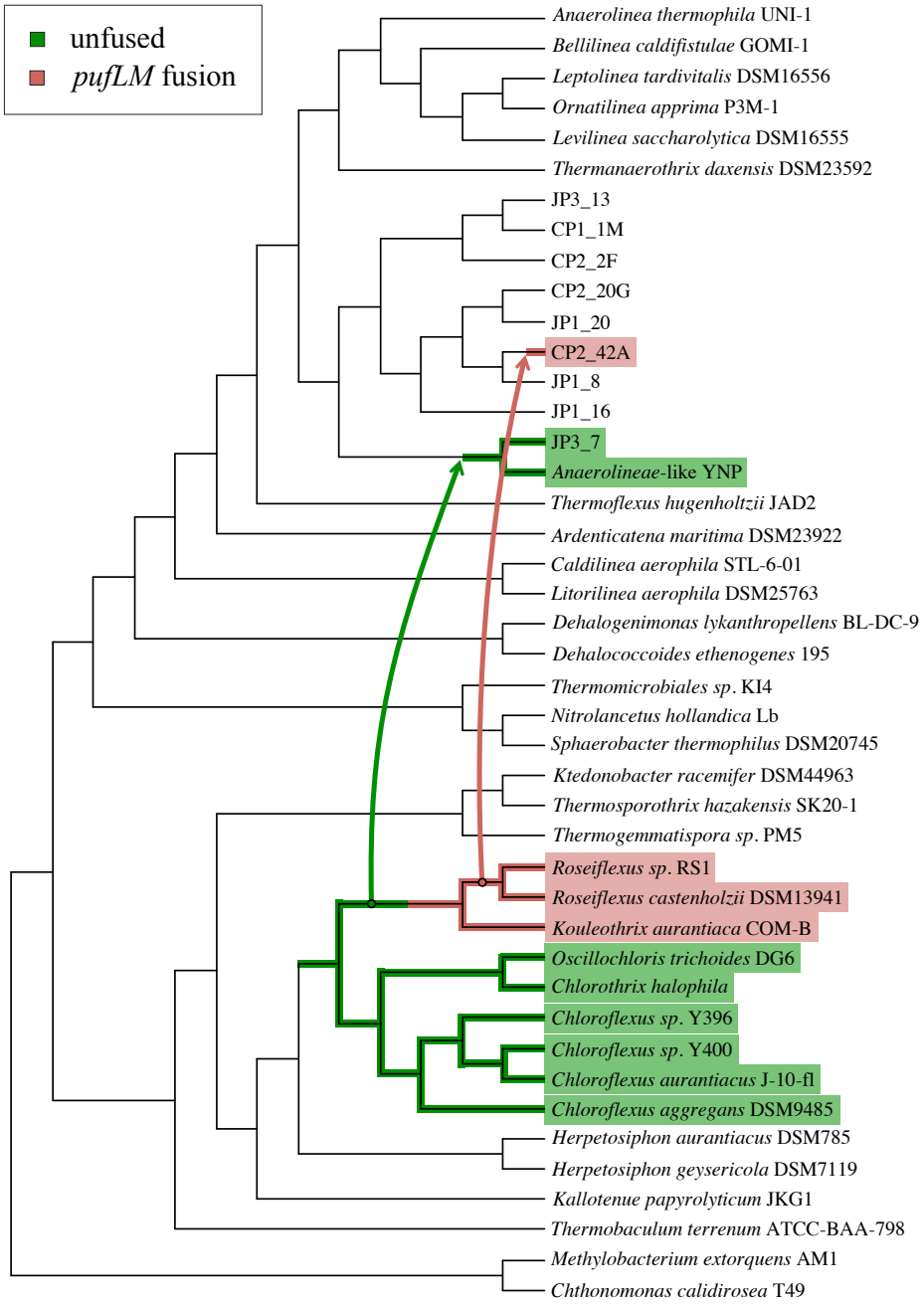


Figure 4: Cartoon of evolutionary scenario of phototrophy in Chloroflexi. Nonphototrophic lineages are in black, lineages with fused PufLM proteins are highlighted in red, and lineages with unfused reaction centers are in green. Arrows represent horizontal gene transfer of phototrophy genes.

	Chloroflexia	Thermomicrobia	Anaerolineae	Caldilineae	Ardenticatenia	Ktedenobacteria	Thermoflexia	Dehalococcoidetes	*Ca.* Thermofonseae
Phototrophy	+ (chlorin-based)	-	-	-	-	-	-	-	Some (chlorin- or rhodopsin-based)
Aerobic respiration	+	+	Genes present	Genes present	+	+	Microaerophilic	Genes rarely present	Genes present
Complex III	bc, ACIII, or both	ACIII or both	ACIII, bc, or neither	bc	bc	bc	bc	Neither	Bc, also ACIII with RCII
Morphology	Filamentous	Rod	Filamentous or rods	Filamentous	Filamentous	Filamentous (branched)	Filamentous	Coccoidal, discs	Unknown
Motility	Gliding	Flagellar	Flagellar or none	-	-	-	-	-	Unknown—no flagellar genes
Other metabolic traits		Nitrite oxidation			Iron and nitrogen respiration			Dehalogenation	Nitrogen respiration
Temperature range	10-67	43-80	20-73	37-65	30-75	17-74	67.5-75	15-35	32-59
% GC	48-62	56-63	48-58	59-65	51.5	54-60	69	49-54	46-63

Table 1: Characteristics of the classes of Chloroflexi, including *Ca*. Thermofonseae described here. The Chloroflexi phylum contains eight recognized classes, with genome sequences available for at least one member of each: Chloroflexia (Garrity & Holt, 2001; Gupta et al., 2013), Thermomicrobia (Hugenholtz&Stackebrandt, 2004, Sorokin et al.

2012), Ktedonobacteria (Cavaletti et al., 2006; Yabe et al., 2010; Chang et al. 2011), Dehalococcoidia (Löffler et al., 2013; Moe et al., 2009), Ardenticatenia (Kawaichi et al., 2013), Anaerolineae and Caldilineae (Yamada et al. 2006), and Thermoflexia (Dodsworth et al. 2014). Table data from this study, Dodsworth et al. 2014, and our previously described Chloroflexi genomes (Ward et al. 2015ab, Hemp et al. 2015abc, Pace et al. 2015).

References:

1. Aziz, RK, D Bartels, AA Best, M DeJongh, T Disz, RA Edwards, ... and O Zagnitko. 2008. The RAST Server: rapid annotations using subsystems technology. BMC genomics 9(1) pp75.
2. Balashov, Sergei P., et al. "Xanthorhodopsin: a proton pump with a light-harvesting carotenoid antenna." *Science* 309.5743 (2005): 2061-2064.
3. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., & Madden T.L. (2008) "BLAST+: architecture and applications." BMC Bioinformatics 10:421.
4. Cavaletti, L., Monciardini, P., Bamonte, R., Schumann, P., Rohde, M., Sosio, M. & Donadio, S. (2006). New lineage of filamentous, spore- forming, gram-positive bacteria from soil. Appl Environ Microbiol 72, 4360–4369.
5. Chang, Yun-Juan, et al. "Non-contiguous finished genome sequence and contextual data of the filamentous soil bacterium Ktedonobacter racemifer type strain (SOSP1-21 T)." *Standards in genomic sciences* 5.1 (2011): 97.

6. Cheng, Q., Day, A., Dowson-Day, M., Shen, G. F. & Dixon, R. The Klebsiella pneumoniae nitrogenase Fe protein gene (nifH) functionally substitutes for the chlL gene in Chlamydomonas reinhardtii. *Biochem. Biophys. Res. Commun.* **329,** 966–975 (2005).

7. Dodsworth, J. A. *et al*. Thermoflexus hugenholtzii gen. nov., sp. nov., a thermophilic, microaerophilic, filamentous bacterium representing a novel class in the Chloroflexi, Thermoflexia classis nov., and description of Thermoflexaceae fam. nov. and Thermoflexales ord. nov. *Int. J. Syst. Evol. Microbiol.* **64,** 2119–2127 (2014).

8. Doolittle, R.F. 1986. Of Urfs and Orfs: Primer on how to analyze derived amino acid sequences, University Science Books, Mill Valley, CA, USA.

9. Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput
Nucleic Acids Res. 32(5):1792-1797

10. Garrity, G. M. & Holt, J. G. (2001). Phylum BVI. Chloroflexi phy. nov. In Bergey's Manual of Systematic Bacteriology, 2nd edn, vol. 1, pp. 427–446. Edited by D. R. Boone, R. W. Castenholz & G. Garrity. New York: Springer.

11. Gouy, M. Guindon, S. & Gascuel., O. (2010) SeaView version 4 : a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evolution 27(2):221-224.

12. Gupta, RS, P Chander, and S George. 2013. Phylogenetic framework and molecular signatures for the class Chloroflexi and its different clades;

proposal for division of the class Chloroflexi class. Nov. into the suborder Chloroflexineae subord. nov., consisting of the emended family Oschillochloridaceae and the family Chloroflexaceae fam. nov., and the suborder Roseiflexineae subord. nov., containing the family Roseiflexaceae fam nov. Antonie van Leeuwenhoek 103 pp99-119.

13. Hanada, S, A Hiraishi, K Shimada, and K Matsuura. 1995. Chloroflexus aggregans sp. nov., a filamentous phototrophic bacterium which forms dense cell aggregates by active gliding movement. IJSEM 45(4) pp676-681.

14. Hanada, Satoshi. "The Phylum Chloroflexi, the Family Chloroflexaceae, and the Related Phototrophic Families Oscillochloridaceae and Roseiflexaceae." *The Prokaryotes*. Springer Berlin Heidelberg, 2014. 515-532.

15. Hansmann, Sabine, and William Martin. "Phylogeny of 33 ribosomal and six other proteins encoded in an ancient gene cluster that is conserved across prokaryotic genomes: influence of excluding poorly alignable sites from analysis." *International Journal of Systematic and Evolutionary Microbiology* 50.4 (2000): 1655-1663.

16. Hemp J, Ward LM, Pace LA, Fischer WW. 2015a. Draft genome sequence of Ornatilinea apprima P3M-1, an anaerobic member of the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01353-15.

17. Hemp J, Ward LM, Pace LA, Fischer WW. 2015b. Draft genome sequence of Ardenticatena maritima 110S, a thermophilic nitrate- and iron-reducing

member of the Chloroflexi class Ardenticatenia. Genome Announc 3(6):e01347-15.

18. Hemp J, Ward LM, Pace LA, Fischer WW. 2015c. Draft genome sequence of Levilinea saccharolytica KIBI-1, a member of the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01357-15.

19. Hug, LA, CJ Castelle, KC Wrighton, BC Thomas, I Sharon, KR Frischkorn, KH Williams, SG Tringe, and JF Banfield. 2013. Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. Microbiome 1 pp1-22.

20. Hug, L. A. et al. A new view of the tree and life's diversity. Nat. Microbiol. (2016). doi:10.1038/nmicrobiol.2016.48

21. Hugenholtz, P. & Stackebrandt, E. (2004). Reclassification of Sphaerobacter thermophilus from the subclass Sphaerobacteridae in the phylum Actinobacteria to the class Thermomicrobia (emended description) in the phylum Chloroflexi (emended description). Int J Syst Evol Microbiol 54, 2049–2051.

22. Igarashi N, Harada J, Nagashima S, Matsuura K, Shimada K, Nagashima KVP. 2001. Horizontal Transfer of the Photosynthesis Gene Cluster and Operon Rearrangement in Purple Bacteria. J. Mol. Evol. 52(4):333–41

23. Inagaki, F, M Suzuki, K Takai, H Oida, T Sakamoto, K Aoki, KH Nealson, and K Horikoshi. 2003. Microbial communities associated with geological

horizons in coastal subseafloor sediments from the Sea of Okhotsk. Appl. Environ. Microbiol. 69(12) pp7224-7235.

24. Kaneoka, Ichiro, Naoki Isshiki, and Shigeo Zashu. "K-Ar ages of the Izu-Bonin islands." *Geochemical Journal* 4.2 (1970): 53-60.

25. Kawaichi, S., Ito, N., Kamikawa, R., Sugawara, T., Yoshida, T. & Sako, Y. (2013). Ardenticatena maritima gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum 'Chloroflexi' isolated from an iron-rich coastal hydrothermal field, and description of Ardenticatenia classis nov. Int J Syst Evol Microbiol 63, 2992–3002.

26. Kawaichi, Satoshi, et al. "Draft genome sequence of a heterotrophic facultative anaerobic thermophilic bacterium, Ardenticatena maritima strain 110ST." *Genome announcements* 3.5 (2015): e01145-15.

27. Klatt, C. G. et al. Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. ISME J. 5, 1262–1278 (2011).

28. Klueglein, N., and A. Kappler. "Abiotic oxidation of Fe (II) by reactive nitrogen species in cultures of the nitrate-reducing Fe (II) oxidizer Acidovorax sp. BoFeN1–questioning the existence of enzymatic Fe (II) oxidation."Geobiology 11.2 (2013): 180-190.

29. Kubo, K, K Knittel, R Amann, M Fukui, and K Matsuura. 2011. Sulfur-metabolizing bacterial populations in microbial mats of the Nakabusa hot spring, Japan. Systematic and Applied Microbiology 34(4) pp293-302.

30. Letunic, Ivica, and Peer Bork. "Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees." Nucleic acids research 44.W1 (2016): W242-W245.

31. Li, D., Luo, R., Löffler, C.M., Leung, C.M., Ting, H.F., Sadakane, K., Yamashita, H. and Lam, T.W., 2016. MEGAHIT v1.0: A Fast and Scalable Metagenome Assembler driven by Advanced Methodologies and Community Practices. Methods.

32. Löffler, F. E., Yan, J., Ritalahti, K. M., Adrian, L., Edwards, E. A., Konstantinidis, K. T., Mu¨ ller, J. A., Fullerton, H., Zinder, S. H. & Spormann, A. M. (2013). Dehalococcoides mccartyi gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, Dehalococcoidia classis nov., order Dehalococcoidales ord. nov. and family Dehalococcoidaceae fam. nov., within the phylum Chloroflexi. Int J Syst Evol Microbiol 63, 625–635.

33. Majumder, Erica LW, Jeremy D. King, and Robert E. Blankenship. "Alternative Complex III from phototrophic bacteria and its electron acceptor auracyanin." *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1827.11 (2013): 1383-1391.

34. McInerney, James O., Mary J. O'connell, and Davide Pisani. "The hybrid nature of the Eukaryota and a consilient view of life on Earth." *Nature Reviews Microbiology* 12.6 (2014): 449-455.

35. Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010) "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1 - 8.

36. Moe, W. M., Yan, J., Nobre, M. F., da Costa, M. S. & Rainey, F. A. (2009). Dehalogenimonas lykanthroporepellens gen. nov., sp. nov., a reductively dehalogenating bacterium isolated from chlorinated solvent-contaminated groundwater. Int J Syst Evol Microbiol 59, 2692–2697.

37. Mulkidjanian, A. Y. *et al*. The cyanobacterial genome core and the origin of photosynthesis. *Proc Natl Acad Sci U S A* **103,** 13126–13131 (2006).

38. Nagashima S, Nagashima KVP. 2013. Comparison of Photosynthesis Gene Clusters Retrieved from Total Genome Sequences of Purple Bacteria. In Genome Evolution of Photosynthetic Bacteria, pp. 151–78. Academic Press

39. Overmann, J. 2008. Green nonsulfur bacteria. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.

40. Oyaizu, H., et al. "The green non-sulfur bacteria: a deep branching in the eubacterial line of descent." *Systematic and applied microbiology* 9.1-2 (1987): 47-53.

41. Pace, Norman R. "Mapping the tree of life: progress and prospects." *Microbiology and Molecular Biology Reviews* 73.4 (2009): 565-576.

42. Pace LA, Hemp J, Ward LM, Fischer WW. 2015. Draft genome of Thermanaerothrix daxensis GNS-1, a thermophilic facultative anaerobe from the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01354-15.

43. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2014. Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research, 25: 1043-1055.

44. Pierson, B. K. & Castenholz, R. W. A Phototrophic Gliding Filamentous Bacterium of Hot Springs , Chloroflexus aurantiacus , gen . and sp. nov. Arch. Microbiol. 100, 5–24 (1974).

45. Raymond, J., Zhaxybayeva, O., Gogarten, J. P., Gerdes, S. Y. & Blankenship, R. E. Whole-Genome Analysis of Photosynthetic Prokaryotes. Science (80-.). 298, (2002).

46. Seviour RJ, Blackall LL. 1999. Current taxonomic status of filamentous bacteria found in activated sludge plants, pp. 122–146. In The Microbiology of Activated Sludge. Springer Netherlands, Dordrecht.

47. Sharma, A. K., Zhaxybayeva, O., Papke, R. T. & Doolittle, W. F. Actinorhodopsins: Proteorhodopsin-like gene sequences found predominantly in non-marine environments. *Environ. Microbiol.* **10,** 1039–1056 (2008).

48. Sorokin, D. Y. *et al*. Nitrification expanded: discovery, physiology and genomics of a nitrite-oxidizing bacterium from the phylum Chloroflexi. *ISME J*. **6,** 2245–2256 (2012).

49. A. Stamatakis: "RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies". In Bioinformatics, 2014

50. Stookey, Lawrence L. "Ferrozine---a new spectrophotometric reagent for iron." Analytical chemistry 42.7 (1970): 779-781.

51. Strous M, Kraft B, Bisdorf R, TegetMeyer H (2012) The binning of metagenomic contigs for microbial physiology of mixed cultures. Frontiers in Microbial Physiology and Metabolism 3:410.

52. Ward LM, Hemp J, Pace LA, Fischer WW. 2015a. Draft genome sequence of Leptolinea tardivitalis YMTK-2, a mesophilic anaerobe from the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01356-15.

53. Ward LM, Hemp J, Pace LA, Fischer WW. 2015b. Draft genome sequence of Herpetosiphon geysericola GC-42, a nonphototrophic member of the Chloroflexi class Chloroflexia. Genome Announc 3(6):e01352-15

54. Ward, LM, A Idai, WW Fischer, and SE McGlynn. 2017a. Microbial diversity and productivity of Jinata Onsen, an iron-rich intertidal hot spring in Japan.

55. Ward, LM, WW Fischer, K Matsuura, and SE McGlynn. 2017b. Cone-forming microbial mats of Nakabusa Onsen, Japan as analogs of Precambrian stromatolites. In prep.

56. Waterhouse, A.M., Procter, J.B., Martin, D.M.A, Clamp, M. and Barton, G. J. (2009) "Jalview Version 2 - a multiple sequence alignment editor and analysis workbench" Bioinformatics25 (9) 1189-1191 doi: 10.1093/bioinformatics/btp033

57. Wätzlich, D. *et al*. Chimeric Nitrogenase-like enzymes of (bacterio)chlorophyll biosynthesis. *J. Biol. Chem*. **284,** 15530–15540 (2009).

58. Williams, T. A., Foster, P. G., Cox, C. J. & Embley, T. M. An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* **504**, 231–236 (2013).

59. Woese CR, Debrunnervossbrinck BA, Oyaizu H, Stackebrandt E, Ludwig W (1985) Gram-positive bacteria—possible photosyn- thetic ancestry. Science 229(4715):762–765

60. Woese, Carl R. "Bacterial evolution." *Microbiological reviews* 51.2 (1987): 221.

61. Xiong, J, WM Fischer, K Inoue, M Nakahara, and CE Bauer. 2000. Molecular evidence for the early evolution of photosynthesis. Science 289 pp1724-1730.

62. Yabe, S., Aiba, Y., Sakai, Y., Hazaka, M. & Yokota, A. (2010). Thermosporothrix hazakensis gen. nov., sp. nov., isolated from compost, description of Thermosporotrichaceae fam. nov. within the class Ktedonobacteria Cavaletti et al. 2007 and emended description of the class Ktedonobacteria. Int J Syst Evol Microbiol 60, 1794– 1801.

63. Yamada M, Zhang H, Hanada S, Nagashima KVP, Shimada K, Matsuura K. (2005). Structural and spectroscopic properties of a reaction center complex from the chlorosome-lacking filamentous anoxygenic phototrophic bacterium *Roseiflexus castenholzii*. *J Bacteriol* **187**: 1702–1709.

64. Yamada, T., Sekiguchi, Y., Hanada, S., Imachi, H., Ohashi, A., Harada, H. &

Kamagata, Y. (2006). Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. Int J Syst Evol Microbiol 56, 1331–1340.

65. Yamada, T and Y Sekiguchi. 2009. Cultivation of uncultured Chloroflexi subphyla: significance and ecophysiology of formerly uncultured Chloroflexi 'Subphylum I' with natural and biotechnological relevance. Microbes Environ. 24pp205–216.

66. Youvan DC, Bylina EJ, Alberti M, Begusch H, Hearst JE. (1984). Nucleotide and deduced polypeptide sequences of the photosynthetic reaction-center, B870 antenna, and flanking polypeptides from *R. capsulata*. *Cell* **37**: 949–957.

67. Yanyushin MF, del Rosario MC, Brune DC, Blankenship RE. 2005. New class of bacterial membrane oxidoreductases. Biochemistry (Mosc.). 44(30):10037–45 Zahnle

68. Zeng, Y., Feng, F., Medova, H., Dean, J. & Kobli ek, M. Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. Proc. Natl. Acad. Sci. 111, 7795–7800 (2014)

Supplementary information:

Kouleothrix aurantiaca

Kouleothrix aurantiaca, a member of the group formerly known as 'Eikenboom morphotype 1851' (Seviour and Burkall 1999), was isolated from activated sludge in an industrial wastewater treatment facility (Kohno et al. 2002). It forms orange-pigmented cells organized into long mm-scale filaments, grows on numerous sugars and pyruvate and by fermentation on certain sugars, and can reduce nitrate to nitrite (Kohno et al. 2002). It is closely related to members of the genus *Roseiflexus* (Beer et al. 2002), however phototrophy has not been observed in these organisms.

The genome of *Kouleothrix aurantiaca* COM-B (JCM 19913) was sequenced as part of a project to expand the phylogenetic breadth of *Chloroflexi* genomes. Genome sequencing was performed at Seqmatic using the Illumina MiSeq sequencing platform. SPAdes 3.1.1 (7) was used to assemble the genome. The genome was screened for contaminants based on sequence coverage, GC composition, and BLAST hits of conserved single copy genes. Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline.

K. aurantiaca falls within the phototrophic Chloroflexia, with a consistent position basal to *Roseiflexus* in both organismal and photosynthetic gene trees (Figures 2 and 3). *K. aurantiaca* encodes for all of the genes required for anoxygenic phototrophy; a type II reaction center (including a fused pufLM and a pufC), a complete bacteriochlorophyll biosynthesis pathway, and a cytochrome bc complex, but no Alternative Complex III. *K. aurantiaca* also encodes for a branched aerobic respiration pathway, including two

Complex Is (NADH dehydrogenase), Complex II (succinate dehydrogenase), two Complex IIIs (cytochrome bc complex), two heme-copper oxygen reductases (A and B-family), and a quinol bd oxidase. In addition it has a NirK nitrite reductase. *K. aurantiaca* encodes a type 1 rubisco and a phosphoribulokinase gene, suggesting that it is capable of carbon fixation via the Calvin Cycle. It does not, however, encode key genes in the 3-hydroxypropionate bicycle used for carbon fixation in *Chloroflexus* and *Roseiflexus* (Klatt et al. 2007).

Supplemental discussion

CP2_42A also encodes a rhodopsin homolog most closely related to xanthorhodopsin, a proton-pumping rhodopsin shown to use light-harvesting antenna carotenoids (Balashov et al. 2005). CP2_42A also encodes for a branched aerobic respiration pathway, including two Complex Is (NADH dehydrogenase), Complex II (succinate dehydrogenase), Complex III (cytochrome bc complex), Alternative Complex III, and two heme-copper oxygen reductases (A and B-family). The CP2_42A genome bin did not recover genes for the Calvin cycle or the 3-hydroxypropionate bicycle, suggesting that it may grow primarily as a photoheterotroph.

JP3_7 also encodes for a branched aerobic respiration pathway, including two Complex Is (NADH dehydrogenase), Complex II (succinate dehydrogenase), Complex III (cytochrome bc complex), two heme-copper oxygen reductases (A and B-family), and a cytochrome bd oxidase.

The bacteriochlorophyll synthesis genes present in JP3_7 and CP2_42A may record a hybrid history of the pathway. While most genes in the pathway (bchX,Y,Z,F,G,P) recovered a close relationship between all of the phototrophic Chloroflexi, the bchI,D, and H (and H-like) genes show a closer relationship between the Chloroflexia and Clorobi to the exclusion of CP2_42A and JP3_7. This may record a hybrid history for bacteriochlorophyll synthesis in the novel Chloroflexi phototrophs, with some genes in the bacteriochlorophyll synthesis pathway being transferred with the reaction center genes from the Chloroflexia, and others deriving from the Chlorobi or other phototrophic bacteria. Alternatively, this may reflect the relative lack of conserved sequence information of short, relatively quickly-evolving proteins relative to the much more conserved and information-rich reaction center and ribosomal proteins.

Phylogenies of electron transport proteins reveal that aerobic respiration using an A family heme copper oxidoreductase and a *bc* complex is a vertically-inherited synapomorphy of the Thermofonseae, while the B family heme copper oxidoreductase and Alternative Complex III found in CP2_42A appear to have been acquired through horizontal gene transfer associated with the acquisition of the type 2 reaction center.

Genes involved in lipopolysaccharide synthesis (e.g. lpxB, lpxC, kdsA) and outer membrane proteins (e.g. bamA) were absent from all Chloroflexi genomes reported here. This is consistent with the proposed single membrane "monoderm" nature of Chloroflexi (Sutcliffe 2010, Sutcliffe 2011) and suggests that this is a conserved feature of the Chloroflexi phylum, though the presence of outer membrane proteins and lipopolysaccharide synthesis in the closely related Armatimonadetes phylum (e.g. Ward et

al. 2017c) suggests that monoderm Chloroflexi are derived from a diderm ancestor and are not representative of an ancestral state.

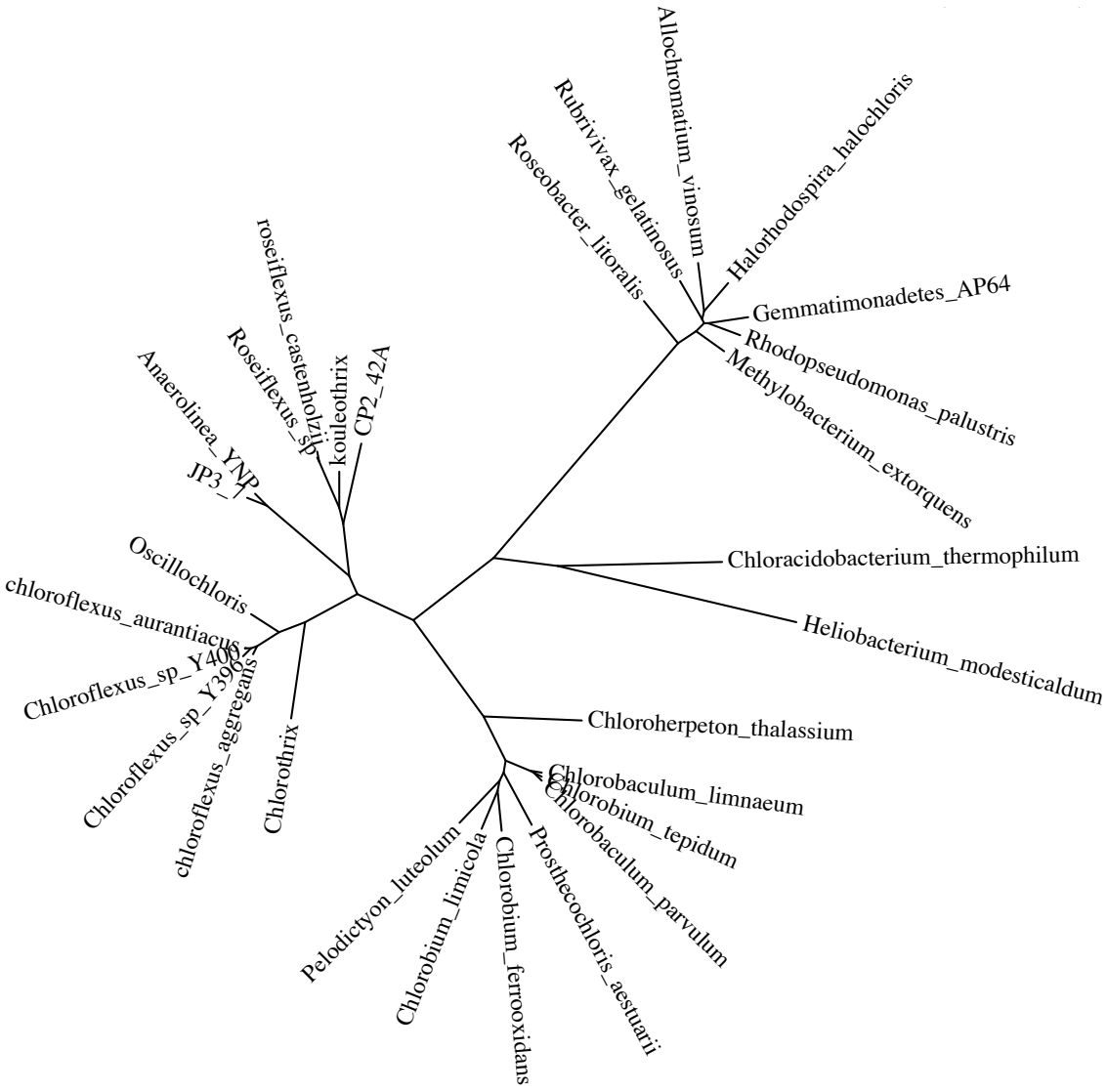
Probability of missing genes

In order to estimate the probability that genes were missing from recovered genome bins, we calculated the probability mass function of recovering zero genes of a particular set from a genome of predicted size, given estimated completeness and assuming random sampling without replacement of individual genes. Though gene size varies significantly and colocalization makes selection of related genes not entirely independent, we assume here that all genes have an equal probability of being selected. This simplifying assumption is reasonable, as recovered phototrophy genes largely reside on separate contigs (suggesting that colocalization is limited) and the length of phototrophy-related genes (e.g. coding for reaction center proteins, bacteriochlorophyll synthases, etc) are within error of average gene length. The calculation took the form of $f(x) = \binom{n}{x}\binom{T-n}{r-x}/\binom{T}{r}$, where f is the probability of recovering x genes of set r from a genome containing T genes of which n are recovered. In the case of our genome bins, n equaled the number of protein coding sequences recovered in each bin, T equaled n divided by the completeness of the genome as estimated by CheckM, and r equaled 6 (representing pufL, pufM, pufC, bchX, bchY, and bchZ). The probability that phototrophy genes existed in in *Ca*. Thermofonseae genomes but was not recovered in our bins ranged from ~0.5 for JP1_191 (at only ~10% completeness) to ~2*10^-13 for JP3_13 (at over 96% completeness). The probability of missing phototrophy genes was only >5% in JP1_191, greatly improving our confidence

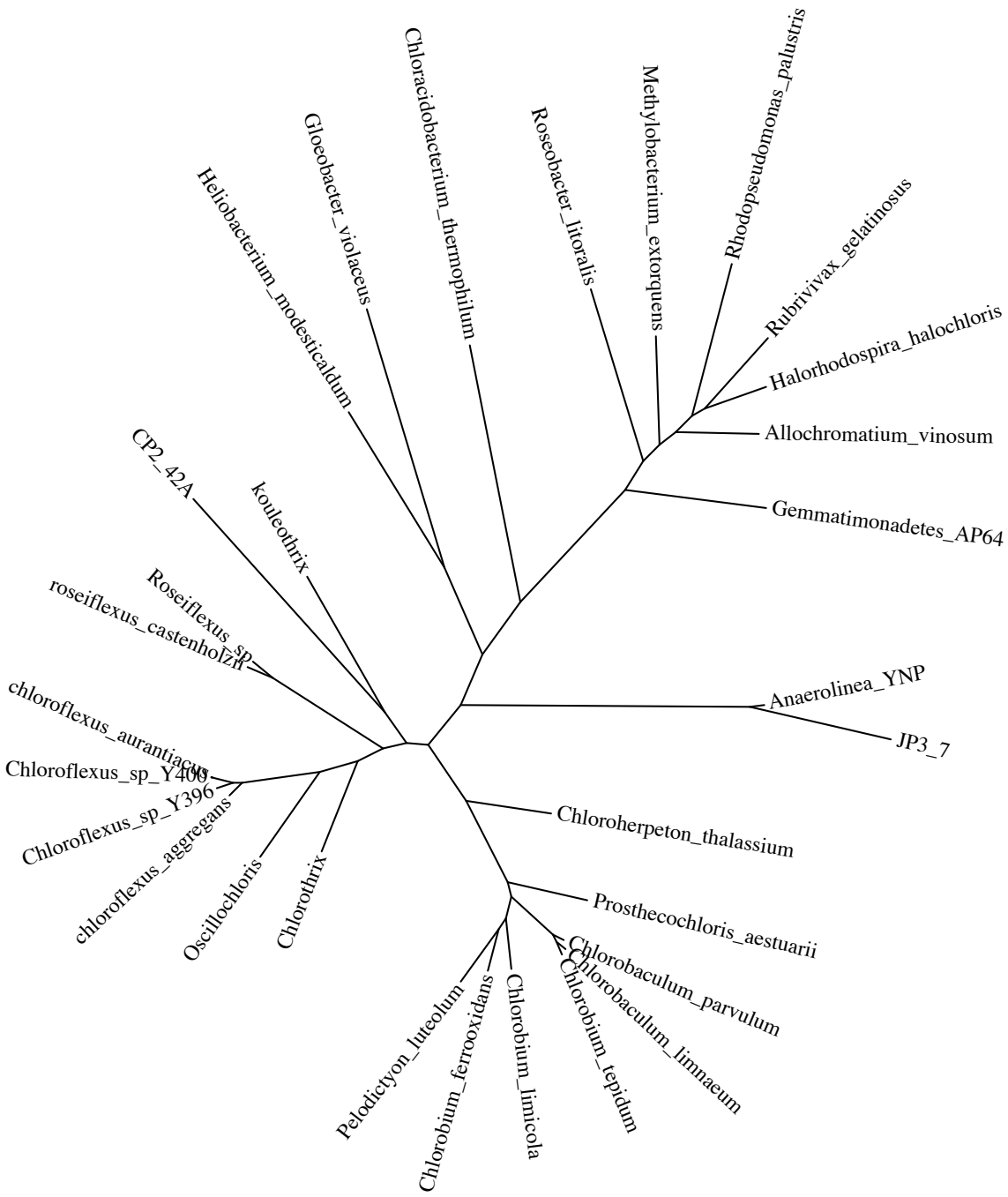
that the absence of phototrophy from most strains of *Ca*. Thermofonseae is a real signal and not an artifact of incomplete genome bins.

A similar calculation can be made for the probability that bchL, N, B, M, or E genes are present in phototrophic Thermofonseae but simply not recovered in our genome bins. The probability of missing all five of these genes is about 0.03% for CP2_42A and less than 0.005% for JP3_7, increasing the possibility that these genes are in fact missing from these genomes, potentially replaced by promiscuous homologs.

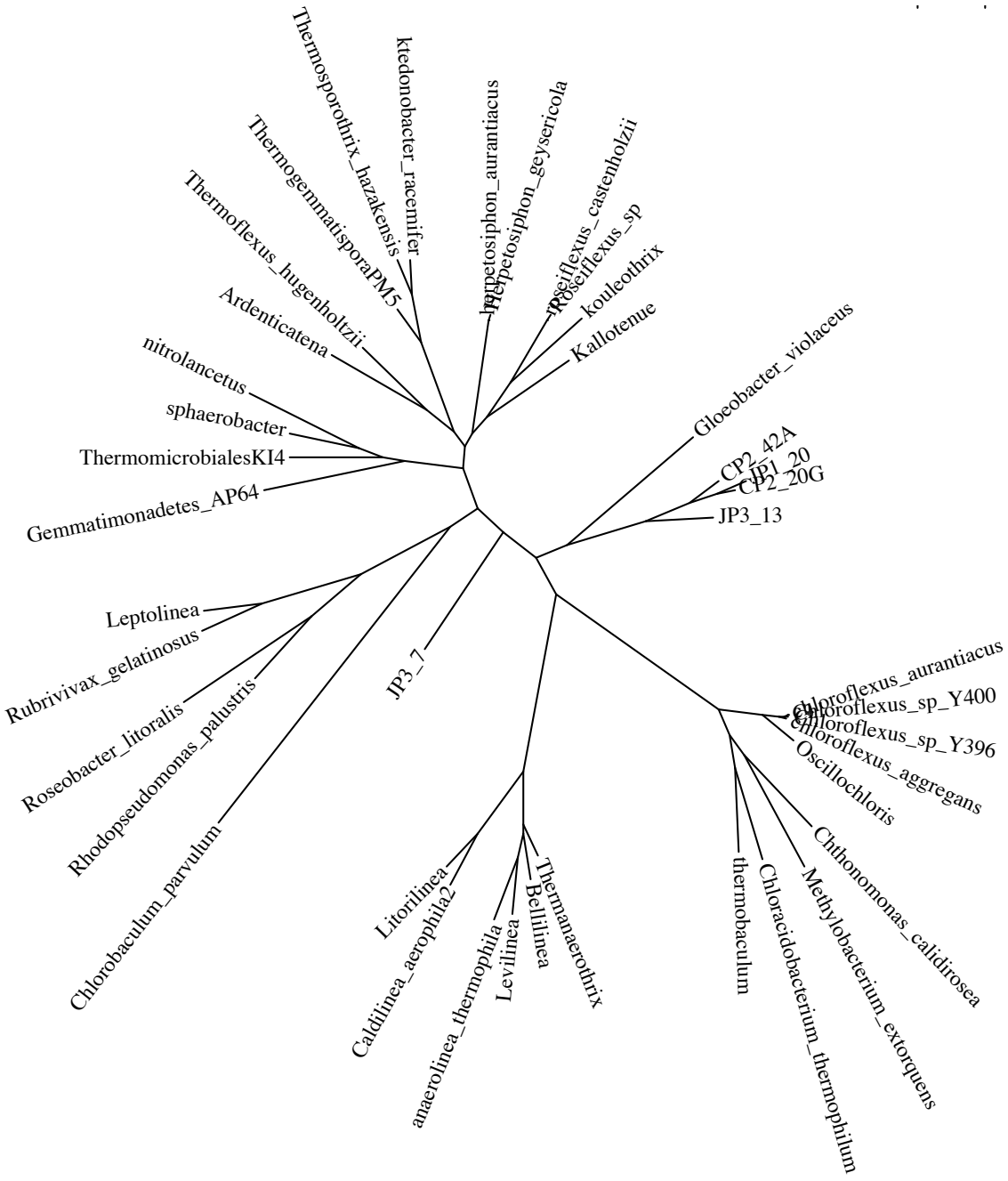
Supplemental figure 1: Unrooted phylogeny of bchXYZ protein sequences



Supplemental figure 2: Unrooted phylogeny of bchIDH protein sequences

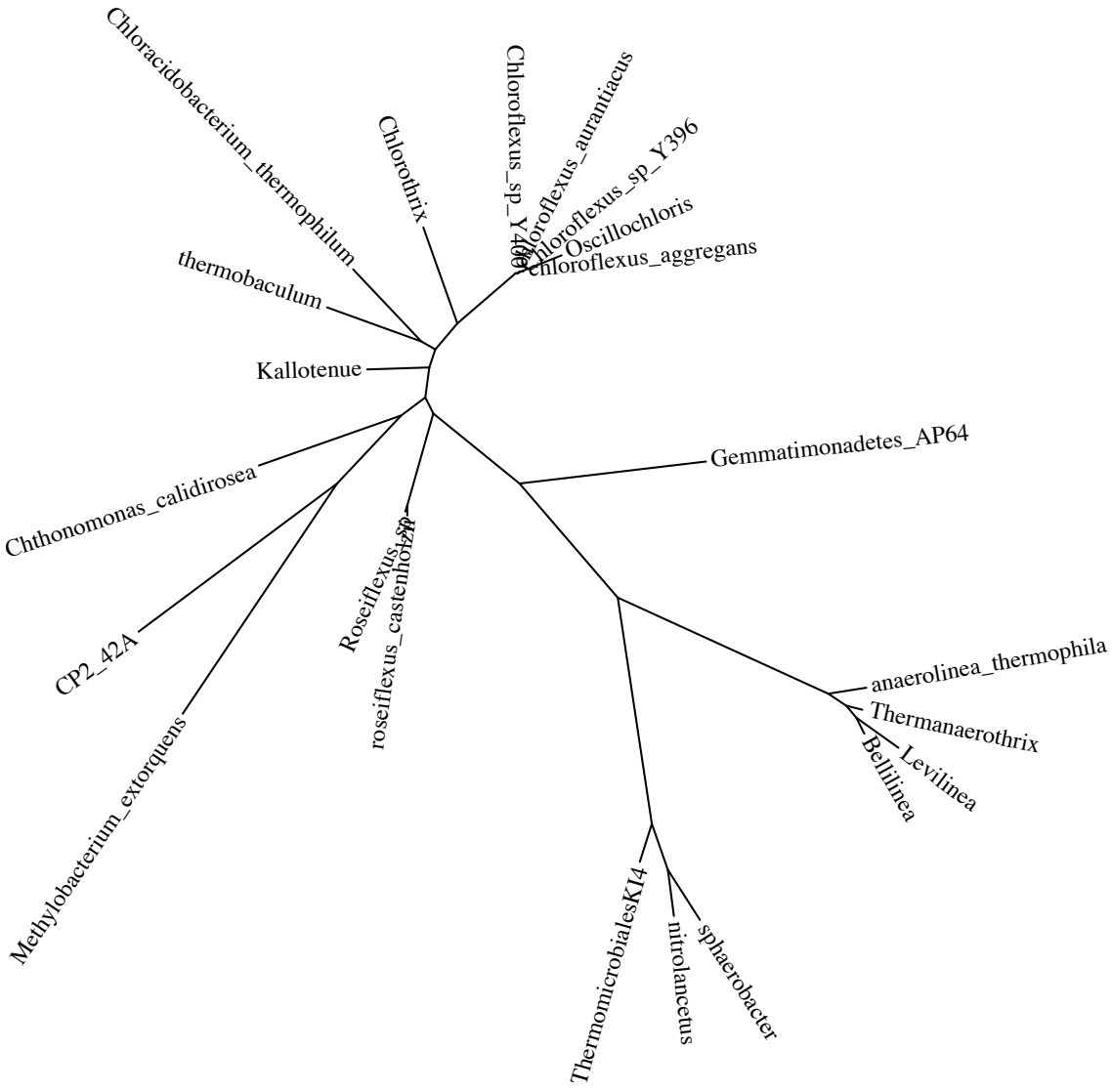


Supplemental figure 3: Unrooted phylogeny of A-family Heme Copper Oxidoreductase protein sequences

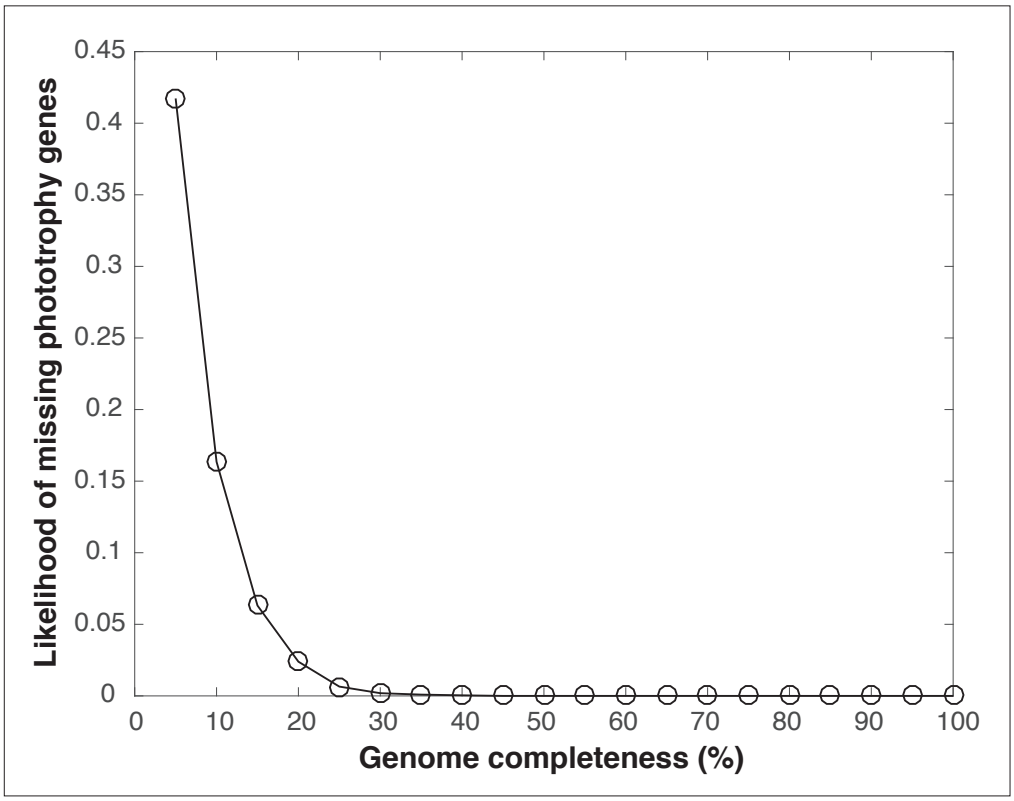


Supplemental figure 4: Unrooted phylogeny of B-family Heme Copper Oxidoreductase protein sequences

Supplemental figure 5: Unrooted phylogeny of Alternative Complex III protein sequences



Supplemental figure 6: Unrooted phylogeny of bc complex protein sequences



Supplemental figure 7: Probability of failure to recover phototrophy genes for a given completeness of genome recovery. Results plotted here are for a simulation following the constraints and logic discussed in the text.

Supplemental Table 1: Genome statistics of sequenced strains

	Genome size	% GC	# coding sequences	# Contigs	Completeness	Contamination	tRNAs	Source
CP1_1M	1.39	59	1182	138	42.28	1.81	14	Nakabusa Cone Pool 1
CP2_2F	1.99	59	1734	20	49.46	0	23	Nakabusa Cone Pool 2
CP2_20G	3.09	48	2678	852	78.54	3.55	32	Nakabusa Cone Pool 2
CP2_42A	3.3	59	2897	2024	79.44	10.42	31	Nakabusa Cone Pool 2
JP1_8	2.21	51	1973	601	58.13	0.13	17	Jinata Pool 1
JP1_16	4.06	44	3238	1764	95.15	17.31	45	Jinata Pool 1
JP1_20	3.36	46	2878	1139	79.09	4.78	34	Jinata Pool 1
JP1_191	0.417	47	334	883	10.63	1.8	7	Jinata Pool 1
JP3_7	3.62	63	3078	1331	87	12.85	46	Jinata Pool 3
JP3_13	3.67	60	3116	1259	96.17	10.87	46	Jinata Pool 3
Kouleothrix aurantiaca	8.7	62	8993	5539	85	0	97	Isolate from wastewater sludge

Supplemental Table 2: Metabolic traits of genomes reported here

	16S	RpoB	RCII	bc	ACII	Rhodopsin	Denitrification	A Fam HCO	B Fam HCO	3HP	Calvin Cycle
CP1_1M	-	+	-	+	-	-	-	-	-	-	-
CP2_2F	-	+	-	-	-	+	-	-	-	-	-
CP2_20G	-	+	-	+	-	-	nirK	+	-	-	-
CP2_42A	-	+	+ (fused)	+	+	+	-	+ (two)	+	-	-
JP1_8	-	+	-	-	-	-	-	-	-	-	-
JP1_16	-	+	-	+	-	-	-	+	-	-	-
JP1_20	-	+	-	+	-	-	nirK, NOR	+ (three)	-	-	-
JP1_191	-	+	-	-	-	-	-	-	-	-	-
JP3_7	-	+	+ (unfused)	+	-	-	-	+	+	-	-
JP3_13	-	+	-	+	-	+	-	+ (two)	-	-	-
Kouleothrix aurantiaca	+	+	+ (fused)	+	-	-	nirK	+	+	-	+

Supplemental references:

1. Beer M, Seviour EM, Kong Y, Cunningham M, Blackall LL, Seviour RJ. 2002. Phylogeny of the filamentous bacterium Eikelboom Type 1851, and design and application of a 16S rRNA targeted oligonucleotide probe for its fluorescence in situ identification in activated sludge. FEMS Microbiol Lett 207:179–183.
2. Klatt, Christian G., Donald A. Bryant, and David M. Ward. "Comparative genomics provides evidence for the 3-hydroxypropionate autotrophic pathway in filamentous anoxygenic phototrophic bacteria and in hot spring microbial mats." *Environmental microbiology* 9.8 (2007): 2067-2078.
3. Kohno T, Sei K, Mori K. 2002. Characterization of type 1851 organism isolated from activated sludge samples. Water Sci Technol 46:111–114.
4. Sutcliffe, I. C. A phylum level perspective on bacterial cell envelope architecture. *Trends Microbiol*. **18,** 464–470 (2010).
5. Sutcliffe, I. C. Cell envelope architecture in the Chloroflexi: A shifting frontline in a phylogenetic turf war. *Environmental Microbiology* **13,** 279–282 (2011).
6. Ward, LM, SE McGlynn, and WW Fischer. 2017c. Draft genomes of two members of the Armatimonadetes phylum recovered from Japanese hot spring metagenomes. In prep.

Chapter 4

MICROBIAL DIVERSITY AND IRON OXIDATION AT OKUOKU-HACHIKUROU ONSEN, A JAPANESE HOT SPRING ANALOG OF PRECAMBRIAN IRON FORMATION

Ward, LM, A Idei, S Terajima, T Kakegawa, WW Fischer, and SE McGlynn. Microbial diversity and iron oxidation at Okuoku-hachikurou Onsen, a Japanese hot spring analog of Precambrian iron formation. Geobiology, in revision.

Abstract:

Banded Iron Formations (BIFs) are rock deposits common in the Archean and Paleoproterozoic (and regionally Neoproterozoic) sedimentary successions. Multiple hypotheses for their deposition exist, principally invoking the precipitation of iron via the metabolic activities of oxygenic, photoferrotrophic, and/or aerobic iron-oxidizing bacteria. Some isolated environments support chemistry and mineralogy analogous to processes involved in BIF deposition, and their study can aid in untangling the factors that lead to iron precipitation. One such process analog system occurs at Okuoku-hachikurou (OHK) Onsen in Akita Prefecture, Japan. OHK is an iron- and CO2-rich, circumneutral hot spring that produces a range of precipitated mineral textures containing fine laminae of aragonite and iron oxides that resemble BIF fabrics. Here, we have performed metagenomic 16S amplicon sequencing of microbial communities across the range of microenvironments in OHK to describe the microbial diversity present and to gain insight into the cycling of iron, oxygen, and carbon in this ecosystem. These analyses suggest that productivity at OHK is based on aerobic iron-oxidizing Gallionellaceae. In contrast to other BIF analog sites,

Cyanobacteria, anoxygenic phototrophs, and iron-reducing microorganisms are present at only low abundances. These observations support a hypothesis where low growth yields and high stoichiometry of iron oxidized per carbon fixed by aerobic iron-oxidizing chemoautotrophs like Gallionellaceae, result in accumulation of iron oxide phases without stoichiometric buildup of organic matter; this system supports little dissimilatory iron reduction, further setting OHK apart from other process analog sites where iron oxidation is primarily driven by phototrophic organisms. This positions OHK as a study area where the controls on primary productivity in iron-rich environments can be elucidated. When compared with geological data, the metabolisms and mineralogy at OHK are most similar to specific BIF occurrences deposited after the Great Oxygenation Event, and generally discordant with those that accumulated before it.

Background:

Banded Iron Formation (BIF) is a characteristic lithotype in many Precambrian basins. These finely laminated, iron-rich (>15% Fe by weight) sedimentary deposits are not only economically critical sources of iron ore (particularly after post-depositional weathering and hydrothermal processes oxidize, leach, and upgrade the ore, Morris 1980, Beukes 1984), but may provide a record of biological activity on the early Earth (e.g. Harder 1919, Konhauser et al. 2002, Kappler et al. 2005).

Despite their prevalence in the early rock record, the processes by which BIFs are deposited are not well understood. It is generally hypothesized that BIFs formed as a result of transport and concentration of ferrous iron (as $Fe^{2+}_{(aq)}$) in seawater under anoxic and sulfur-poor conditions, followed by oxidation and precipitation of iron as ferric iron phases

(Holland 1973, Drever 1974, Holland 1984). However, it is important to note that the primary mineralogy that made up the precursor sediments to BIF remains uncertain in many cases, and the significance of iron oxidation during BIF deposition remains an area of active debate (Konhauser et al. 2002, Kappler et al. 2004, Kappler et al. 2005, Bekker et al. 2013, Posth et al. 2013, Rasmussen et al. 2013, Rasmussen et al. 2015, Rasmussen et al. 2016, Tosca et al. 2016). Different hypotheses for BIF deposition invoke a range of iron oxidation processes, including abiotic iron oxidation by UV light (e.g. Cairns-Smith 1978, Francois 1986), indirectly biologically by O_2 sourced from oxygenic photosynthesis by Cyanobacteria (e.g. Cloud 1973), or directly biologically by aerobic iron-oxidizing bacteria (e.g. Chan et al. 2016) or by anaerobic iron-oxidizing phototrophic bacteria (e.g. Widdel et al. 1993, Kappler et al. 2005). BIFs do not form in marine sedimentary basins today due to the O_2 and sulfate content of seawater (Canfield 1998), which prevents the mobilization and concentration of sufficient amounts of dissolved iron in seawater and shallow pore fluids (Holland 1973). However, a range of potential analog environments can be observed today that may reveal key processes associated with the deposition of BIFs. These include permanently stratified lakes, such as Lake Matano in Indonesia (Crowe et al. 2008), and iron-rich hot springs, such as Chocolate Pots in Yellowstone National Park (Pierson et al. 1999). These systems contain anoxic, iron-rich waters that produce iron oxides and other Fe-bearing phases through a range of processes. In Lake Matano, iron oxide production is thought to be driven largely by photoferrotrophy (Crowe et al. 2008), while at Chocolate Pots hot spring iron oxidation is driven primarily by abiotic reaction of iron with O_2 produced by Cyanobacteria (Pierson et al. 1999, Pierson and Parenteau 2000, Trouwborst

et al. 2007). Terrestrial hot springs can provide unique environments in which high concentrations of dissolved iron can exist at circumneutral conditions due to rock weathering by anoxic source waters; they also provide access to novel microbial diversity which is rare or absent from typical surface environments. While these hot spring fluids are not perfect compositional mimics of Precambrian seawater, the geomicrobiological processes they support and the resulting facies and fabrics can provide process analogs for understanding depositional mechanics of at least some BIFs. This is particularly useful for evaluating the relative roles that different microbiological processes (e.g. photoferrotrophy, aerobic iron oxidation, and oxygenic photosynthesis) may have played in the deposition of different BIF facies. A crucial but understudied aspect of BIF deposition in modern analog environments is the relative delivery fluxes of iron oxides and organic carbon to sediments, and the average oxidation state of the resultant iron formations; trends in the proportion of ferrous to ferric iron in ancient BIFs have long been observed in the rock record (e.g. Klein 2005)—and these can reflect the redox state of the environment at the time of deposition, as well as the particular physiologies and metabolisms responsible for iron oxidation (Fischer and Knoll 2009)(Supplemental Information). Though it has been little exploited, the metabolic yield and efficiency of the microbial metabolisms driving iron oxidation will be a factor in the mineralogy and redox chemistry of diagenetically stabilized iron formation lithologies. Currently, knowledge of the microbial metabolisms and processes that lead to different mineral assemblages (and ferric:ferrous ratios) in modern environments remains meager. Characterization of the microbial communities in the context of the organic carbon content and redox state of iron in the solids accumulating in BIF analogs deposited

differentially by photoferrotrophs, Cyanobacteria, and aerobic iron oxidizers will help to connect observations of ancient BIFs to the geobiological processes responsible for iron oxide deposition

In this study, we investigated a novel BIF-analog environment at Okuoku-hachikurou Onsen (OHK) in Akita Prefecture, Japan. OHK is an iron- and carbonate-rich, circumneutral hot spring with anoxic source waters, which produces extensive iron oxide and aragonite travertine with mineralogical and textural features resembling BIF (Takashima et al. 2011). To study the microbial diversity of OHK and evaluate the potential roles of different microbial metabolisms in producing mineral deposits at OHK, we collected mineral precipitates and filtered spring water from points along the outflow of the stream for analysis using microscopy and characterization of the microbial community via metagenomic 16S amplicon sequencing. These data reveal that productivity and iron oxidation at OHK is primarily driven by aerobic, iron-oxidizing taxa related to the genus *Sideroxydans*, and consequently the environment produces sediments that are organic-lean and contain a high proportion of ferric iron phases—an early diagenetic precondition to the highly oxidized BIF facies deposited in certain environments after the Great Oxygenation Event.

Materials and Methods:

Geological context and sedimentology of OHK:

OHK is located in Akita Prefecture, Japan, at 40.407925N, 140.754744E (Figure 1), in an active region of the Tohoku volcanic arc, generated by the subduction of the Pacific Plate; the local bedrock geology consists of Miocene–Holocene green tuff and

felsic volcanic rocks (Shimazu et al. 1965). OHK consists of a single subsurface water source that originates from a mining exploration borehole drilled in the 1960s. This borehole emerges into a 2 m diameter source pool, with a submerged shelf around the edge of the borehole at 1.2 m depth. At the time of our study in November 2015, water emerged from the borehole at 44.3°C, high in Fe(II), very low in oxygen (Table 1), and displayed continuous and vigorous ebullition of CO_2 (Figure 2). The source pool contains abundant suspended fine flocs of hydrous iron oxides. The early crystalline iron oxides phases at OHK consist of ferrihydrite (Takashima et al. 2011), though this may age into more ordered iron oxide minerals such as hematite, goethite, or lepidocrosite in travertines. From the source pool to the rotenburo (Figure 2), the stream flows over a mineralized substrate comprised of well-lithified, finely laminated aragonite and iron oxide travertine with a smooth texture and mottled coloration ranging from white aragonite to orange and red iron oxides (Figure 2). Surfaces of pools, channels, and canals are coated in fine, orange colored iron oxide flocs. Mature iron oxide phases within travertine appear dark red. Travertine accumulates around pool edges as lobate walls ~10 cm in width. Where spring water overflows the edges of pools travertine terraces develop with a characteristic ~5 cm step size and up to about a half meter in total height. These terraces superficially resemble those found in both circumneutral- and acidic-iron-rich systems elsewhere (e.g. Yellowstone National Park, Fouke 2011, and the Tintillo River, Spain, España et al. 2007); the occurrence of similar large scale morphologies across very geochemically different environments suggests that hydrology—as much as geochemistry or microbial processes—plays a role in the travertine morphologies (Fouke 2011).

An old outflow channel emerges from the source pool, but was largely inactive at the time of our sampling; flow into this channel has been blocked and redirected through canals to the rotenburo. This channel is ~5 cm deep, partially mixes with the source pool due to swashing caused by CO_2 ebullition, but is otherwise stagnant. Thin plates of aragonite (<1 mm) coat the water surface of the channel, and in some areas thin microbial biofilms have developed on the bottom of the channel.

Water emerging from the source pool predominantly flows through shallow (5-10 cm) canals into a ~1 m diameter pool used for bathing (rotenburo). Rock surfaces in the canals and rotenburo are coated in iron oxide flocs. Water flows out from the rotenburo through additional shallow canals and ultimately develops into a sheeted flow on the hill slope, where continual degassing of CO_2 leads to formation of bubbles. Many bubbles become encrusted with aragonite and mineralized (Figure 2).

While no substantial accumulation of microbial biomass is visible within the source pool, rotenburo, or canals, thin (<1 mm) patchy green biofilms occur along the rim of the source pool and old stream (Figure 2). These were sampled for the Shallow Source and Old Stream Mineral samples.

Downstream, OHK preserves unique mineralized bubbles. As CO_2 degasses from the spring water, dissolved inorganic carbon concentrations drop and the carbonate saturation state increases, prompting aragonite precipitation. This occurs sufficiently quickly in OHK that bubbles develop coatings of aragonite and are mineralized *in situ*. These bubbles encrust and provide a possible preservation mechanism for a wide range of organic materials, including leaves, arthropods, and biofilms in this section of the spring.

Preservation of organic structures by iron minerals has been observed previously in acidic iron-rich environments at Rio Tinto, Spain, where fossil structures endure for millions of years (Fernandez-Remolar and Knoll 2008). The rapid mineralization of iron oxides and aragonite at OHK may therefore also serve as a mechanism for preserving key biological features of this environment over geologic timescales.

Sample collection:

Samples were collected from 5 sites at OHK (Figure 2): Deep Source (1.2 m deep in the source pool at the borehole), Shallow Source (surface water and a thin biofilm along the edge of the source pool), Old Stream (water and thin biofilm along the semi-stagnant blocked outflow), Canal (along the flow from the source pool to the rotenburo), and Bubble Pool (downstream in a 15 cm deep pool coated in mineralized bubbles).

From, each site, both a "Mineral" and a "Water" sample were collected for 16S amplicon sequencing, targeting surface-associated and pelagic microbial communities, respectively. "Mineral" samples consisted of scrapings of thin biofilms, mineral precipitates, or whole mineralized bubbles from surfaces of travertines or cobbles in the bottom of water channels. Mineral samples were collected using sterile forceps and spatulas (~0.25 cm^3 of material). "Water" samples consisted of cells and sediment filtered from water using sterile syringes and 0.2 micron Sansyo (Sansyo Co., Tokyo, JP) filters (~50-200 ml of water filtered, until filter began to clog). Cells were lysed and DNA preserved in the field using Zymo Terralyzer BashingBead Matrix and Xpedition Lysis Buffer (Zymo Research, Irvine, CA). Cells were disrupted immediately by attaching tubes to the blade of a cordless reciprocating saw (Black & Decker, Towson, MD) and operating

for 1 minute. Samples for geochemical analysis consisted of water collected via sterile syringe and filtered immediately through a 0.2 micron Sansyo filter.

Geochemical analysis:

Dissolved oxygen (DO), pH, and temperature measurements were performed *in situ* using an Extech DO700 8-in-1 Portable Dissolved Oxygen Meter (FLIR Commercial Systems, Inc., Nashua, NH). Iron concentrations were measured using the ferrozine assay (Stookey 1970) following acidification with 40 mM sulfamic acid to inhibit iron oxidation by O_2 or oxidized nitrogen species (Klueglein and Kappler 2013). Ammonia/ammonium concentrations were measured using a TetraTest NH_3/NH_4^+ Kit (TetraPond, Blacksburg, VA) following manufacturers instructions but with colorimetry of samples and NH_4Cl standards quantified with a Thermo Scientific Nanodrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA) at 700 nm to improve sensitivity and accuracy. Anion concentrations were measured via ion chromatography on a Shimadzu Ion Chromatograph (Shimadzu Corp., Kyoto, JP) equipped with a Shodex SI-90 4E anion column (Showa Denko, Tokyo, JP).

Total Organic Carbon (TOC) contents were assessed for filtered water and mineral precipitates from the source pool, as well as mineral precipitates near the canal. For dissolved organic carbon measurements, 3 liters of spring water was filtered using 0.45 micron Sansyo glass fiber filters. For mineral precipitates, approximately 300 g samples were collected. Carbonate carbon was removed via dissolution with HCl, and residues were folded into tin capsules and analyzed for carbon content via elemental analyzer (Thermoscientific Flash 2000) with a detection limit of 0.005% C by weight.

Samples of sediment from the source pool were characterized via SEM-EDS (SU5500; Hitachi, Tokyo, JP) and μM X-ray diffraction by TEM (JEM-3010; JEOL, Tokyo, JP).

Sequencing and analysis:

Following return to the lab, DNA was extracted and purified with a Zymo Soil/Fecal DNA extraction kit. The V4-V5 region of the 16S rRNA gene was amplified from each extract as well as negative controls using archaeal and bacterial primers 515F (GTGCCAGCMGCCGCGGTAA) and 926R (CCGYCAATTYMTTTRAGTTT) (Caporaso et al., 2012). DNA was quantified with a Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA) according to manufacturer's instructions following DNA extraction and PCR steps. All samples yielded PCR amplicons when viewed on a gel after initial pre-barcoding PCR (30 cycles). Duplicate PCR reactions were pooled and reconditioned for five cycles with barcoded primers. Samples for sequencing were submitted to Laragen (Culver City, CA) for analysis on an Illumnia MiSeq platform. Sequence data were processed using QIIME version 1.8.0 (Caporaso et al., 2010). Raw sequence pairs were joined and quality-trimmed using the default parameters in QIIME. Sequences were clustered into de novo operational taxonomic units (OTUs) with 99% similarity using UCLUST open reference clustering protocol (Edgar, 2010). Then, the most abundant sequence was chosen as representative for each *de novo* OTU (Wang et al., 2007). Taxonomic identification for each representative sequence was assigned using the Silva-115 database (Quast et al., 2013) clustered at separately at 99% and at 97% similarity. Singletons and contaminants (OTUs appearing in the negative control datasets)

were removed. 16S sequences were aligned using MAFFT (Katoh et al. 2002) and a phylogeny constructed using FastTree (Price et al. 2010). Alpha diversity was estimated using the Shannon Index (Shannon 1948) and Inverse Simpson metric (1/D) (Simpson 1949; Hill 1973). All statistics were calculated using scripts in QIIME and are reported at the 99% and 97% OTU similarity levels. Multidimensional scaling (MDS) analyses and plots to evaluate the similarity between different samples and OHK environments were produced in R using the vegan and ggplot2 packages (R Core Team 2014, Oksanen et al. 2016, Wickham 2009).

Results and Discussion:

Geochemistry:

Geochemistry measurements of OHK source water are summarized in Table 1, while geochemical gradients along the stream outflow are summarized in Figure 3. Water emerging from the source was 44.3°C, very low in dissolved oxygen (<15 μM), had a pH 6.8, and contained substantial concentrations of dissolved iron (114 μM Fe^{2+}), and 22 μM NH_3/NH_4^+. After emerging from the borehole, the spring water rapidly exchanges gases with the air due to turbulent mixing associated with water flow and CO_2 ebullition, and DO rose to 30 μM in the shallow source pool. Dissolved organic carbon in the source pool was below the limit of detection (0.005% C by weight). Organic carbon content of mineral precipitates at the source pool was 0.01% by weight, rising to 0.02% near the canal. SEM-EDS analysis of source pool precipitates detected Fe and O as the major elements, and selected area X-ray diffraction by TEM revealed amorphous Fe-hydroxides as the solid phase with no detection of hematite or other crystalline phases (Figure 4). As water flows

downstream from the source pool, it cools slightly, degases CO_2, and continues to absorb atmospheric O_2. By the time water reaches the bubble pool, the spring water is 40°C, pH 7.5, and contains 7.7 μM NH_3/NH_4^+, 109 μM O_2 and no detectable dissolved iron. Concentrations of major anions were largely stable along the spring outflow, with 1350 mg/l Cl^-, 620 mg/l SO_4^{2-}, and no detectable NO_2^-, NO_3^-, or HPO_4^-.

Like many modern BIF analog hot spring sites, under the environmental conditions at OHK, abiotic iron oxidation is expected to proceed spontaneously. Following rate equations for abiotic iron oxidation from Singer and Stumm (1970), the rate of abiotic iron oxidation proceeds proportionately to concentrations of dissolved Fe^{2+}, O_2, and the activity of OH^-, of the form $-d[Fe^{2+}]/dt=k[Fe^{2+}]P_{O2}[OH^-]^2$, where k is a rate constant equal to about $8x10^{13}$ $liter^2$ $mole^{-2}$ atm^{-1} min^{-1}at 25°C, and increasing approximately 10-fold for every 15°C of temperature increase (Stumm and Lee 1961). In the deep source pool, where $[O_2]$ was below the detection limit of 0.5 mg/L (~15 μM), abiotic iron oxidation rates are therefore expected to be less than ~4.35 μM/min. This expected rate would be expected to increase between ~8.7 μM/min in the shallow source pool to ~12.7 μM/min in the canal, the last sampling location at which dissolved iron was detectable. These values are significant, and as a result abiotic iron oxidation likely contributes substantially to the production of iron oxides at OHK. Nonetheless, the abundance of iron oxidizing bacteria and morphology of iron oxide precipitates (below) demonstrates that there is a viable and active niche for biological iron oxide production. This is further supported by estimates of flow rates through OHK correlated against changes in iron oxide abundance from the edge of the Shallow Source Pool to the Canal, where flow rates of ~0.33 m/s through canals with

cross sections of ~10 cm by 5 cm suggest flow rates of ~1.67 L/s. Over distances of 2 m from the Shallow Source to the Canal, [Fe^{2+}] declines from ~114 μM to ~83 μM, or a change of 31 μM in approximately 6 seconds; this is equivalent to about 100-fold faster than expected for purely abiotic rates. Even assuming substantial backflow and mixing, this leaves a substantial role for aerobic iron oxidizing bacteria in explaining the decline in iron concentrations along the hot spring outflow. These results are also consistent with estimates from Kusama and Murakami (2001) that suggest aerobic iron oxidizing bacteria can increase iron oxidation rates by up to 4 orders of magnitude above those expected from purely abiotic reactions.

Recovered microbial diversity:

In total, we recovered 141,125 sequences from the 10 samples at OHK (Table 2). Reads per sample ranged from 2,176 for the Canal Mineral sample to 27,454 reads for the Old Stream Mineral sample (median 15,247, mean 14,112, and standard deviation 8,654). With the exception of the Old Stream Mineral sample, water samples consistently recovered more sequence reads than mineral samples (mean of 5,926 versus 18,983). Assessment of sampling depth was estimated using Good's Coverage (Good 1953). On average, 90% of the microbial community was recovered from OHK samples at the 99% OTU level (ranging from 75% coverage in the Canal Mineral sample to 96% in the Old Stream Mineral sample) and 94% at the 97% OTU level (83% for the Canal Mineral sample to 98% for the Old Stream sample).

Most of the taxonomic and abundance variation was related to location along the outflow and therefore the local environmental and redox conditions and consequent metabolic opportunities (Figure 3, Figure 5). The community composition of the water samples all appeared relatively similar to each other, with low dissimilarity between water samples and mineral samples from the source pool. Downstream, mineral samples appear dissimilar both from water samples and from each other (Figure 5).

Relative abundances of microbial taxa as revealed by 16S surveys can be useful for predicting metabolisms driving geochemical cycles and producing the mineral deposits observed at OHK (Table 3 and Figure 3). In particular, the contributions of various iron oxidizers and phototrophs to primary productivity along the spring path (Figure 3) can be estimated due to these metabolisms being fairly well conserved within bacterial taxa (e.g. Emerson et al. 2010, Chan et al. 2016). Analysis of the most abundant taxa at OHK revealed significant roles for organisms with diverse iron metabolisms, including aerobic and phototrophic iron oxidation, as well as trace iron reducers. While the relative contribution of abiotic iron oxidation at OHK is not constrained, sequence data suggest that aerobic iron oxidizing bacteria are the dominant biological drivers of iron oxidation in the hot spring. Also present were diverse phototrophs associated with both oxygenic and anoxygenic phototrophic clades; these organisms were observed at only trace abundance in most samples, but anoxygenic phototrophs were enriched in biofilms from the Old Stream, the rim of the Source Pool, while Cyanobacteria were abundant in the Bubble Pool. A small but diverse assortment of microbes associated with nitrogen cycling was present at

OHK, despite low concentrations of nitrogen compounds in the spring water (Supplementary Information)

Aerobic iron-oxidizing bacteria:

Among the most abundant taxa at OHK are members of the betaproteobacterial family Gallionellaceae, most likely members or close relatives of the aerobic iron-oxidizing genera *Gallionella* or *Sideroxydans*. *Gallionella* and *Sideroxydans* are neutrophilic iron oxidizers that use molecular oxygen to oxidize dissolved Fe(II) to Fe(III) oxides while conserving energy for growth and autotrophic carbon fixation (Kucera and Wolfe 1957, Emerson and Moyer 1997, Emerson et al. 2013); they are commonly found in terrestrial iron-rich systems (Emerson et al. 2010). Aerobic iron-oxidizing bacteria have characteristically low growth yields (Neubauer et al. 2002) due to the modest potentials of Fe(II)/Fe(III) redox couples and resulting requirement for reverse electron transfer to achieve the sufficiently low potential electrons needed for carbon fixation (Bird et al. 2011).

Gallionellaceae are abundant in both the mineral and water fractions of the Deep Source (37 and 18%, respectively), but otherwise appear to be dominantly associated with the water fraction of other samples (28% average in water samples, 8.5% average in mineral samples)(Table 3). This suggests that these organisms are predominantly planktonic or associated with suspended sediment particles rather than forming biofilms on stationary mineral surfaces. Based on the sequence data, members of the Gallionellaceae appear to be the first iron oxidizers and primary producers to act on the upwelling spring water as it mixes with atmospheric O_2, driving the bulk of early biological iron oxidation at

OHK and producing much of the iron oxide sediment that is transported along the spring outflow. Gallionellaceae were fairly diverse, including 169 OTUs at the 97% identity cutoff; however, the two most abundant OTUs were both ~97% similar to *Sideroxydans lithotrophicus* ES-1, and represented more than 92% of the total Gallionellaceae sequences at OHK. The abundance of these two OTUs drives the overall trend in iron oxidizer abundance (Figure 3).

Also present, albeit at lower abundance (up to ~1.5% relative abundance)(Table 3), are members of the zetaproteobacterial family Mariprofundaceae, another group of neutrophilic iron oxidizers. Iron-oxidizing Zetaproteobacteria are more commonly found in marine settings, particularly in deep ocean basins associated with hydrothermal iron sources (Emerson et al. 2010). Despite a similar physiology, Mariprofundaceae are not closely related to Gallionellaceae or other aerobic iron oxidizers, instead forming a distinct class within the Proteobacteria (Emerson et al. 2007). Zetaproteobacteria have previously been identified in relatively saline terrestrial iron- and CO_2-rich systems (e.g. Emerson et al. 2016), sometimes co-occurring with *Gallionella* (Crossey et al. 2016); the discovery of the co-occurrence of these organisms at OHK provides further support to the overlapping ecological niches of these classically "marine" and "terrestrial" iron oxidizing bacteria.

Members of the Gallionellaceae are typically associated with cold iron-oxidizing environments, and not hot springs (Delong et al. 2014), while members of the Mariprofundaceae have been observed to have an upper growth temperature of 30°C (Emerson et al. 2010). OHK, with source water temperatures ~44°C may therefore support unique thermotolerant strains of these bacteria.

Members of the family Comamonadaceae were also fairly abundant (~1-12%) in OHK samples (Table 3). This family of Betaproteobacteria includes members such as *Acidovorax ebreus,* a nitrate-reducing anaerobic iron oxidizer (Byrne-Bailey et al. 2010), as well as iron reducers such as *Rhodoferax ferrireducens* (Finneran et al. 2003) and the iron-oxidizing bacterium *Leptothrix* (van Veen et al. 1978). However, the taxonomic affinity of the Comamonadaceae at OHK is insufficiently resolved to confidently assess the contribution of this group to iron cycling in this environment.

Electron microscopy of mineral precipitates from the source pool and canal revealed alternating laminations of aragonite-rich and iron-oxide-rich material; imaging of iron oxide bands following dissolution of carbonates with HCl revealed that iron oxides were made up of a mixture of amorphous and sheath-like tubular structures (Figure 4). While iron oxide sheaths are typically associated with the betaproteobacterial iron-oxidizing genus *Leptothrix* (Emerson et al. 2010), they can also be produced by diverse iron oxidizers including some strains of Zetaproteobacteria (Fleming et al. 2013). Though we tentatively regard these mineralized filaments as biological in origin, it is unclear what organisms are responsible for the production of iron oxide sheaths observed at OHK. Sub-micron, amorphous particulate iron oxides are characteristic of iron oxidation by *Sideroxydans* (Emerson and Moyer 1997), and so the prevalence of this iron-oxide morphology is consistent with this genus of the Gallionellaceae being major contributors to iron oxidation at OHK. The abundance of amorphous iron oxide particles also supports the assignment of Gallionellaceae OTUs at OHK to *Sideroxydans* rather than stalk-forming *Gallionella* (Emerson et al. 2010).

Cyanobacteria:

Cyanobacteria were abundant in the Bubble Pool Water sample, where they made up ~37% of all sequence reads, but were of much lower abundance in samples collected upstream (Table 3 and Figure 3). Though Cyanobacteria are sometimes underrepresented in iTag datasets as a result of poor DNA yield or amplification biases (e.g. Parada et al. 2015, Trembath-Reichert et al. 2016), the paucity of Cyanobacteria in upstream OHK samples was confirmed by fluorescent microscopy, in which cells displaying cyanobacterial autofluorescence were observed abundantly in samples from the downstream Bubble Pool samples but not in the Source Pool (Supplementary Figure 2).

It has been demonstrated that in iron-rich systems where Cyanobacteria are abundant and productive, only ~1% of O_2 released oxidizes ferrous iron, with the remainder escaping to the atmosphere (Rantamäki et al. 2016). Thus, given the inefficiency of cyanobacterial oxygen fluxes for oxidizing dissolved iron, and the scarcity of Cyanobacteria upstream at OHK where iron oxidation is taking place, Cyanobacteria do not appear to be major contributors to iron oxidation at OHK.

Cyanobacteria present are predominantly members of Subsection III, Family I. This group includes *Leptolyngbya*, a genus of filamentous non–heterocystous Cyanobacteria that has appeared in clone libraries from OHK (Takashima et al. 2011) and is common in other hot springs of similar temperatures (e.g. Roeselers et al. 2007, Bosak et al. 2012).

Members of deeply branching nonphototrophic Cyanobacteria clades are a minor but notable component of OHK samples (up to 0.9% abundance). While the Cyanobacteria phylum has traditionally been considered to exclusively contain oxygenic phototrophs,

several deep-branching nonphototrophic clades have recently been described within the Cyanobacteria phylum, including Melainabacteria, a sister group to oxygenic Cyanobacteria (i.e. Oxyphotobacteria), as well as deeper-branching clades (Ley et al. 2005, Di Rienzi et al. 2013, Johnson et al. 2013b, Soo et al. 2014, Soo et al. 2015). These deep-branching Cyanobacteria—particularly the clades SHA-109 and ML635J-21, which branch basal to all other Cyanobacteria—are thought to be ancestrally nonphototrophic, and can help to better constrain the evolutionary history of Cyanobacteria and therefore oxygenic photosynthesis (e.g. Shih et al. 2016, Fischer et al. 2016, Soo et al. 2017). These clades are found at higher abundance at OHK than most other environments, and OHK could provide a valuable resource for investigating members of this understudied group via metagenomic sequencing, incubations, or isolation.

Anoxygenic phototrophs and relatives:

Members of several taxa made up of or containing anoxygenic phototrophs were present at low abundance in OHK samples (Table 3). These include the Rhodospirillalles, Rhodobacteraceae, Rhodocyclaceae, Chloroflexaceae, Chlorobiales, and Chromateaceae. Some of these taxa (e.g. Chloroflexaceae, Chlorobiales) are made up almost exclusively of phototrophs, while others (e.g. Rhodobacteraceae) contain members with a wide diversity of metabolisms, only some of which are phototrophic (Overmann and Garcia-Pichel 2013, Fischer et al. 2016). Shotgun metagenomic or culture-based analysis will be necessary to confirm whether the members of these taxa present at OHK are phototrophic. Though these taxa individually represent no more than a few percent of the sequence reads at any given site (overall average ~0.88% of reads per taxon), this population is quite diverse and in sum

represents a sizable fraction of the total microbial community (3-20%)(Figure 3). Putative anoxygenic phototrophs are most abundant in the Shallow Source and Old Stream mineral samples (13.8% and 20.4% of total abundance, respectively), dropping to ~3% in the Deep Source sample. Even in the downstream Bubble Pool samples, anoxygenic phototrophs make up ~5% of the total abundance of sequences. Sequences associated with anoxygenic phototrophs were more abundant in mineral samples than in water samples (~9% vs. ~5%), suggesting that these organisms grow preferentially attached to solid surfaces rather than planktonically.

All samples contained relatively abundant sequence reads belonging to the Chlorobi phylum (Table 3). The Chlorobi are classically known as the Green Sulfur Bacteria due to the anaerobic sulfur-oxidizing anoxygenic phototrophic lifestyle of its earliest described members (Davenport et al. 2010, Bryant and Liu 2013). This includes iron-oxidizing anoxygenic phototrophs such as *Chlorobium ferrooxidans* (Heising et al. 1999), which employ a metabolism thought to be relevant to Archean BIF deposition (e.g. Kappler et al. 2005). Due to increased environmental sequencing and new isolation efforts, however, the Chlorobi phylum is now known to also contain aerobic photoheterotrophs (Liu et al. 2012b, Stamps et al. 2014) and nonphototrophs (Podosokorskaya et al. 2012). The majority of Chlorobi sequences found in OHK appear to fall within the Chlorobi order Ignavibacteria, a basal clade of Chlorobi whose known members include versatile heterotrophic metabolisms but no known phototrophy pathways (Iino et al. 2010, Liu et al. 2012a). It therefore appears that photoferrotrophy by Chlorobi is not driving iron oxidation at OHK, though metagenomic sequencing and assembly of OHK Chlorobi genomes will be

necessary to confirm that phototrophy is not present in these organisms. Ignavibacteria appear to be a common component of hot spring microbial communities: these organisms were first isolated from a Japanese hot spring (Iino et al. 2010), and are found at high abundance in Chocolate Pots hot springs in Yellowstone National Park (Fortney et al. 2016).

The Ignavibacteria found at OHK had only low similarity to described strains, with more than 50% of Ignavibacteria reads (primarily from the Deep Source and Canal Water samples) from an OTU ~91% similar to *Melioribacter roseus* P3M, a moderately thermophilic facultative anaerobe (Kadnikov et al. 2013). Approximately 20% of Ignavibacteria reads (primarily from the Deep Source Water and Old Stream Mineral samples) were 93% similar to *Ignavibacteria album* JCM 16511.

Members of the bacterial phylum Chloroflexi were remarkably abundant in OHK samples (Table 3). The Chloroflexi were classically been described as the Green Nonsulfur Bacteria due to the anoxygenic phototrophic metabolism of their earliest described members (Overmann 2008), but it is now recognized that the phylum is much more genetically and metabolically diverse (Yamada and Sekiguchi 2009). Metabolic characters in the Chloroflexi largely follow class-level taxonomic patterns but with a number of notable exceptions, such as the nonphototrophic predatory *Herpetosiphon* within the predominantly phototrophic Chloroflexia class (Kiss et al. 2011, Ward et al. 2015b). The most abundant Chloroflexi at OHK belong to the class Anaerolineae, which were abundant in all samples (up to ~11%). The Anaerolineae have generally been isolated as obligately anaerobic heterotrophs (e.g. Sekiguchi et al. 2003, Yamada et al. 2006), but genome

sequencing has revealed the capacity for aerobic respiration in diverse members of this clade (e.g. Hemp et al. 2015a, Hemp et al. 2015b, Pace et al. 2015, Ward et al. 2015a). Furthermore, a genome for an organism closely related to the Anaerolineae with genes for photosynthesis has been assembled from a Yellowstone National Park metagenome (Klatt et al. 2011). It is therefore unclear what metabolisms may be present in Anaerolineae at OHK, and isolation or metagenomic sequencing of these organisms will be necessary to determine what role they may be playing in this environment. The Anaerolineae at OHK were very diverse, including 480 OTUs at the 97% cutoff. This included a distinct population at the Deep Source (both Water and Mineral Samples), the Shallow Source Mineral sample, and more downstream samples. The three most abundant Anaerolineae OTUs were most closely related to *Thermomarilinea lacunifontana* (84-88% similarity); *T. lacunifontana* is an anaerobic heterotroph isolated from a shallow hydrothermal system in Japan (Nonoura et al. 2013). The most abundant Anaerolineae OTU in the Shallow Source Water sample (making up ~7% of all Anaerolineae at OHK) was 89% identical to *Ornatilinea apprima*. The most abundant OTU in the Old Stream Mineral sample was 91% identical to *Longilinea arvoryzae*. Both *O. apprima* and *L. arvoryzae* are described as obligately anaerobic fermenters capable of degrading sugars and proteins (Yamada et al. 2007, Podosokorskaya et al. 2013).

Conclusions:

OHK is a unique ecosystem supporting novel microbial communities as well as serving as an intriguing process analog for Precambrian banded iron formation deposits. Future activity measurements of community members, for example by metagenomics and

stable isotope probing will be necessary to further define microbial activities in this system. Based on microscopy and 16S amplicon data, the microbial communities at OHK appear to be supported primarily by aerobic iron oxidation occurring in and near the source pool. In the old stream, mineral attached anoxygenic phototrophs become more significant, while Cyanobacteria become abundant only in the most downstream samples. This predominance of lithoautotrophs over phototrophs is rare at the Earth's surface today, and provides a contrast to other modern BIF analog sites. For instance, in Lake Matano iron oxidation is thought to be driven largely by photoferrotrophs (Crowe et al. 2008). At Chocolate Pots hot spring in Yellowstone National Park—perhaps the most geochemically-similar system to OHK that has been extensively studied—iron oxidation is thought to be primarily driven abiotically by O_2 produced *in situ* by cyanobacterial mats (Trouwborst et al. 2007). Furthermore, relative to Chocolate Pots, OHK supports very little in the way of well-developed microbial mats, with only thin, patchy biofilms.

The absence of substantial microbial biomass accumulation near the source pool at OHK can be considered the result of two separate phenomena: 1) the paucity of photosynthetic Cyanobacteria, and 2) the poor growth yields of the iron-oxidizing bacteria that do occur. Neither of these issues is fully resolved, but hypotheses for their causes can be made based on the geochemistry and mineralogy of OHK and related systems, as well as aspects of the physiologies driving iron oxidation.

It has been proposed that high iron concentrations are toxic to Cyanobacteria, and that this may have played a role in delaying the oxygenation of the Archean atmosphere (Swanner et al. 2015). In principle, ferrous iron toxicity may help explain the absence of

Cyanobacteria in OHK until the most downstream samples, where most iron has already been oxidized and precipitated. The absence of Cyanobacteria at OHK is challenging to explain, however, as other iron-rich systems (e.g. Chocolate Pots Hot Spring, Trouwborst et al. 2007) support productive cyanobacterial populations, and in other systems oxygenic photosynthesis and aerobic iron oxidation have been shown to co-occur (Hegler et al. 2012, Mori et al. 2015). At Fuschna Spring in the Swiss Alps, *Gallionella*-dominated communities occur in the iron-rich, low-oxygen, high-flow conditions within the flow channel, but Cyanobacteria- and iron reducer-rich microbial mats accumulate along the edges where flow is less pronounced (Hegler et al. 2012). This suggests that flow regime may also play a role in determining the microbial community of iron-rich systems; OHK experiences high flow rates, and turbulent mixing in the source pool, and this may play a role in limiting the development of phototrophic microbial mats. Turbulence may inhibit the development of phototrophic microbial mats, while simultaneously being advantageous to aerobic iron oxidizing bacteria by helping them shed accumulated iron oxides, limiting encrustation by their metabolic byproducts.

A reasonable hypothesis for the poor development of biofilms by aerobic iron oxidizing bacteria at OHK could be related to the low growth yield of aerobic iron-oxidizing microbes relative to phototrophs. Based on electron balance for an average oxidation state of zero for organic carbon, the maximum possible efficiency of autotrophic iron oxidation is 1 mole of CO_2 fixed for every 4 moles of Fe(II) oxidized. However, measured yields of aerobic iron oxidizers are typically much lower, on the order of 1 mole of CO_2 fixed for every 40 moles of Fe(II) oxidized (Neubauer et al. 2002). Yields for

photoferrotrophs appear to be much more efficient, approximating the ideal stoichiometry of 4 Fe: 1 C (Ehrenreich and Widdel 1994). While electron transfer in neutrophilic iron oxidizers has not been extensively characterized, this difference in yield appears to fundamentally come down to the redox potential of iron oxidation reactions, which (while quite variable depending on environmental pH and mineralogy of iron oxides) are too electrochemically positive to directly reduce $NAD(P)^+$ and therefore be useful for carbon fixation (Bird et al. 2011). In order to grow, these organisms must consume proton motive force (PMF) to run electrons "uphill" to lower redox potentials in order to generate the NAD(P)H needed to reduce CO_2 (Bird et al. 2011). In aerobic iron oxidizers, this requires large fluxes of iron oxidation to maintain sufficient PMF to fix carbon, while phototrophic iron oxidizers can rely on cyclic electron flow through their reaction centers to build PMF sufficiently to allow stoichiometric iron oxidation and carbon fixation. This relatively poor growth yield of aerobic iron oxidizing bacteria at OHK results in organic carbon-lean mineral precipitates (<0.02% organic carbon by weight), in contrast to phototroph-dominated iron-rich systems like Chocolate Pots hot spring where organic carbon contents can be in excess of 1% organic carbon by weight (Parenteau and Cady 2010). Ultimately it is the overproduction of ferric iron relative to carbon fixation during aerobic iron oxidation that sets the budgets for subsequent carbon and iron cycling in the environment.

Significantly, 16S sequence reads associated with environmentally common iron-reducing microbes (e.g. Shewanellaceae, Geobacteraceae) occur at only very low abundances at OHK, with a maximum abundance of the Deltaproteobacteria family Geobacteraceae of 0.79% in the Deep Source Water sample (Table 3). This is in contrast to

other iron rich, neutral pH systems including hot and cold springs and groundwater seeps (e.g. Hegler et al. 2012, Blöthe and Roden 2009, Roden et al. 2012, Fortney et al. 2016). This may partially reflect the relative fluxes of iron oxidation versus organic carbon fixation at OHK; *i.e*. there is insufficient organic carbon being fixed in this environment to fuel substantial iron oxide respiration. Meanwhile, turbulent mixing ensures that oxygen-poor microenvironments do not develop. Since molecular oxygen is available in substantial excess of organic carbon, heterotrophy at OHK never depletes O_2 sufficiently to make iron oxides a favorable electron acceptor. In contrast, in systems with more substantially developed microbial mats or less efficient mixing, oxygen can become depleted deeper in mat fabrics or in diffusion-limited boundary layers, driving local anoxia and the shift toward iron respiration. The lack of substantial iron reduction at OHK is consistent with the predominance of ferric iron minerals in the OHK deposits (Takashima et al. 2011)(Figure 3).

The relative paucity of organic carbon, the dominantly ferric iron content of sedimentary laminations, and the primary role of aerobic iron oxidation together make OHK most similar to Proterozoic-type BIFs, deposited after the GOE. BIF composition varies through time, with BIFs of different ages likely forming via different processes. This variability is likely driven by changes in primary productivity and remineralization pathways caused by the evolution of oxygenic photosynthesis in Cyanobacteria and the subsequent oxygenation of the atmosphere at the GOE *ca*. 2.3 Ga. While substantial debate exists about the antiquity of Cyanobacteria, multiple lines of evidence suggest that oxygenic photosynthesis evolved only shortly before the GOE (Johnson et al. 2013a, Ward

et al. 2016, Fischer et al. 2016, Shih et al. 2016), and therefore molecular oxygen derived from photosynthesis was unlikely to play a role in Archean BIF formation. The hypothesis that the deposition of these BIFs occurred via phototrophic iron oxidation is consistent with the predominantly ferrous composition of Archean and early Paleoproterozoic iron formations (Fischer and Knoll 2009)(Supplementary Information), and is discordant with mechanisms relying on aerobic iron oxidation; photoferrotrophy results in stoichiometric amounts of iron oxide and organic carbon delivered to sediments, which promotes an environment conducive to substantial amounts of iron reduction during burial and diagenesis (Konhauser et al. 2005, Johnson et al. 2007, Li et al. 2013). After the GOE, molecular oxygen was sufficiently abundant in the atmosphere Earth surface environments that it could be used to drive aerobic iron oxidation. Additionally, it has been observed that the organic carbon content of BIF tends to be inversely correlated to the proportion of residual ferric iron, with the most iron oxide-dominated BIFs containing on the order of 0.01% or less organic carbon by weight (Klein 2005, Fischer et al. 2014). Proterozoic BIFs, such as the syn-glacial iron formations that co-occur with Snowball Earth episodes late in Neoproterozoic (Cryogenian) time (Kirschvink 1992, Hoffman et al. 1998), are dominantly comprised of ferric iron phases. This is consistent with expectations for their deposition via aerobic iron oxidation (i.e. excess deposition of iron oxides relative to organic carbon, limiting subsequent iron reduction), and most closely resembles the mineralogy of the materials currently being deposited at OHK. By contrasting OHK with other BIF process analog sites where iron oxidation predominantly occurs by different processes (e.g. phototrophy), it may be possible to open a window into the ecology, mineralogy,

productivity, and other aspects of BIF deposition across the GOE, with OHK representing an endmember in which the iron oxide component of BIF sediment was deposited primarily by aerobic iron oxidation.

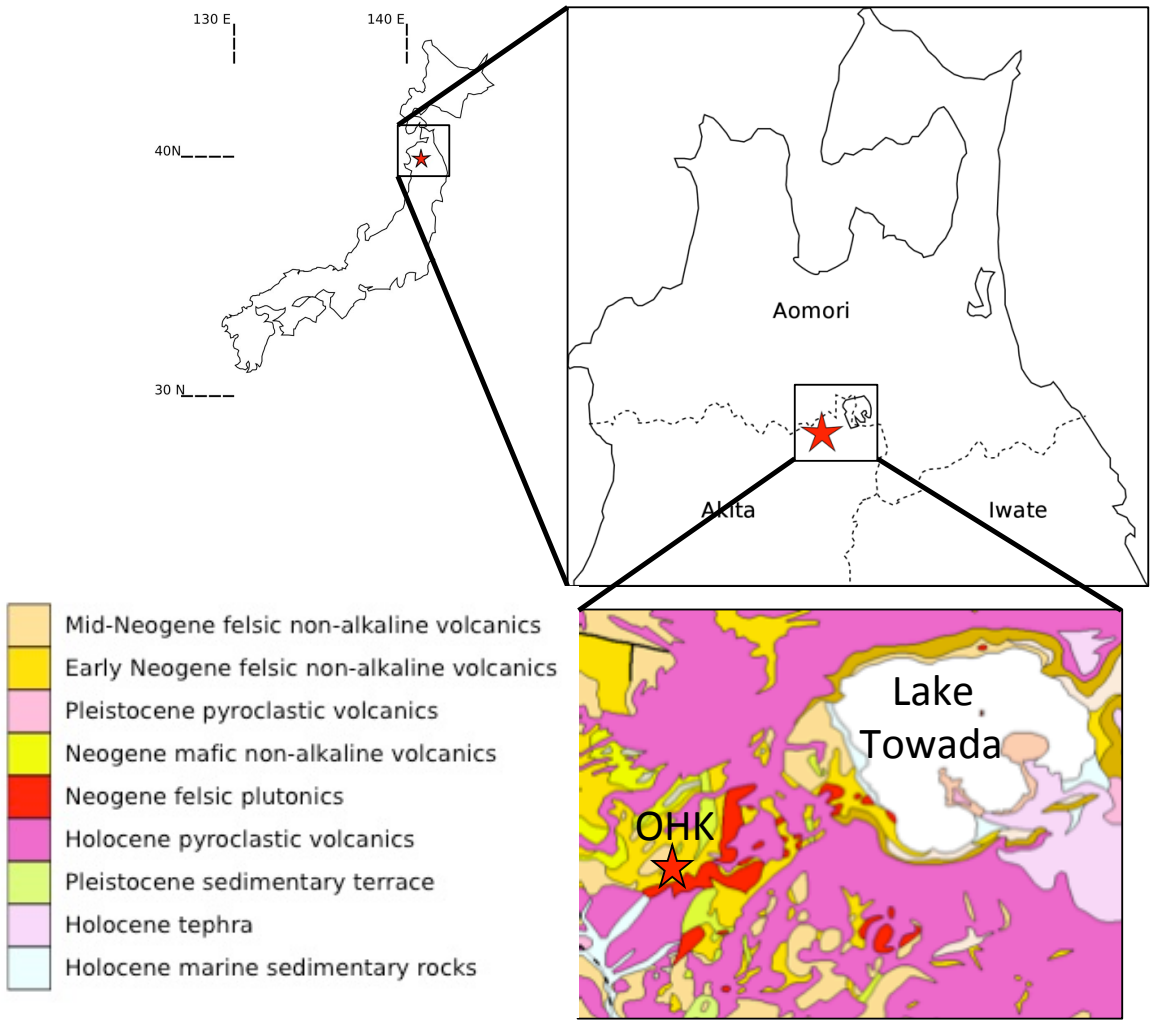


Figure 1: Location of Okuoku-hachikurou Onsen (OHK) in Akita Prefecture, Japan. Inset highlights the local geology of the Lake Towada region, modified from Geological Survey of Japan 2012. The bedrock geology at OHK consists predominantly of felsic volcanics.

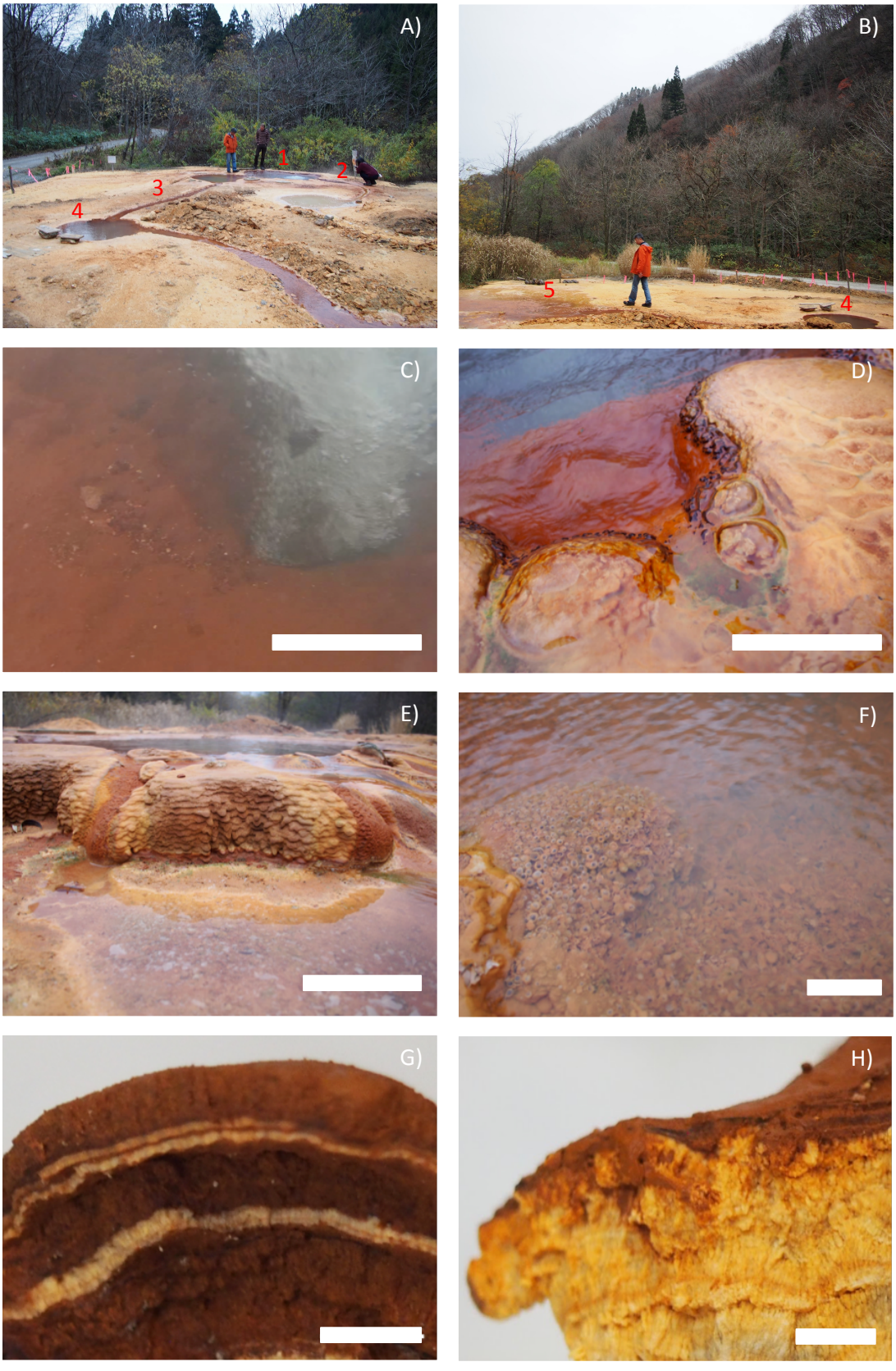

A)
1
2
3
4
B)
5
4
C)
D)
E)
F)
G)
H)

Figure 2: Context photographs of OHK. A) Photograph of OHK facing north toward source pool. Source pool (1), old stream (2), canal (3) and rotenburo (4) visible. People for scale (~1.8 m). B) Image of OHK facing southwest toward rotenburo and bubble pool. Water flowing from the rotenburo (4) develops into a sheeted flow which spreads across the hillside, filling the bubble pool (5) where degasing CO_2 bubbles become encrusted in aragonite. Person for scale (~1.8 m). C) Underwater photograph of source. Water and CO_2 bubbles emerge from the borehole at right. Scale bar is 10 cm. D) Lobate wall of travertine at the rim of the source pool, with green-pigmented biofilms (sampled as Shallow Source Mineral sample) visible. Scale bar is 10 cm. E) Close up of travertine on the west edge of the source pool. Scale bar is 50 cm. F) Close up photograph of the bubble pool, showing aragonite-mineralized CO_2 bubbles. Scale bar 5 cm. G,H) Close ups of ferrihydrite and aragonite laminations in travertine. Red layers are primarily composed of ferrihydrite, while white layers are predominantly aragonite. Scale bars shown are 5 mm.

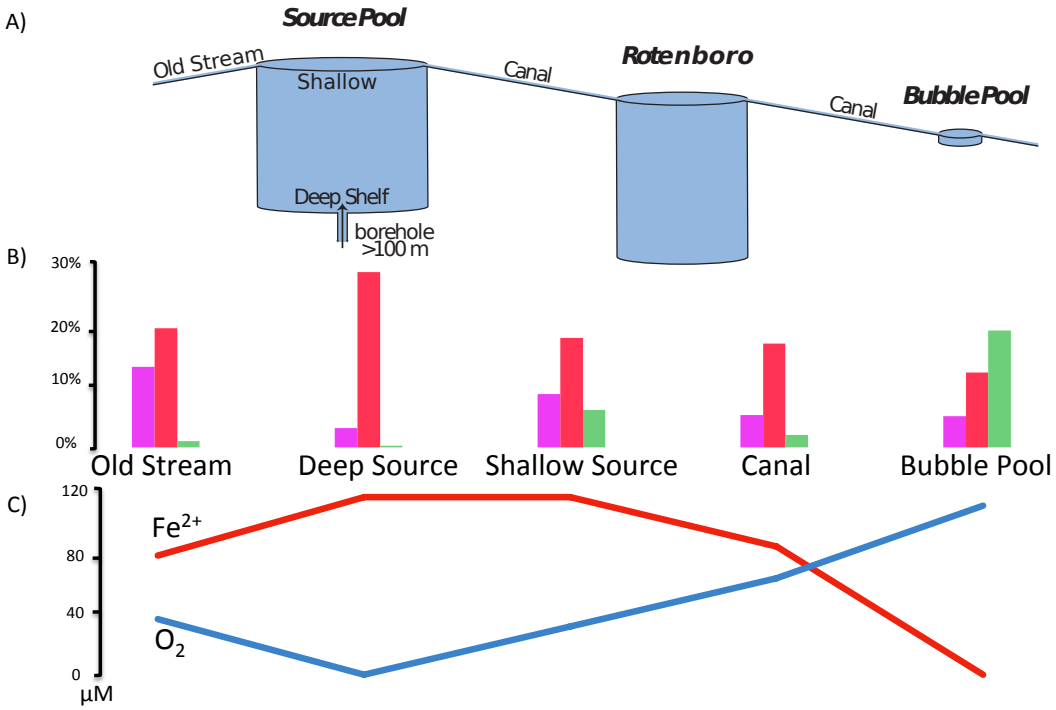


Figure 3: Summary of the geochemistry and iron-oxidizing community composition of OHK ordinated along the flow paths in the system. A) Schematic of OHK flow paths, keyed to the sample names. Mineral and Water samples were collected from Old Stream, Deep Source, Shallow Source, Canal, and Bubble Pool localities. B) Relative abundance of phototrophic and iron-oxidizing taxa along the flow path. Data here include taxa listed in Table 5 and discussed in the text with confidently assigned metabolisms, but not including taxa such as Anaerolineae whose metabolisms cannot be predicted with available taxonomic resolution. Relative abundance of taxa have been summed and then averaged per site between Mineral and Water samples. Aerobic iron oxidizers are shown in red, anoxygenic phototrophs in purple, and oxygenic Cyanobacteria in green. Aerobic iron oxidizers dominate near the source pool, while anoxygenic phototrophs are abundant in the Old Stream and Shallow Source Mineral samples, and oxygenic Cyanobacteria are highly

abundant only in the Bubble Pool Water sample. C) Concentrations of dissolved O_2 and Fe^{2+}. Source waters are depleted in O_2 but high in dissolved iron, while downstream Fe^{2+} concentrations drop and O_2 increases. The majority of iron oxidation appears to occur under microoxic conditions where aerobic iron oxidizers are most prevalent, while anoxygenic phototrophs are largely restricted to biofilms, and Cyanobacteria are only significant after the iron has already been removed from solution. Changes in temperature (44.3-40.7 °C) and concentrations of chloride (~37-39 mM) and sulfate (~6.3-6.65 mM) along the hot spring outflow were minor and not expected to contribute significantly to changes in microbial community or mineral precipitates, and so are not displayed here.

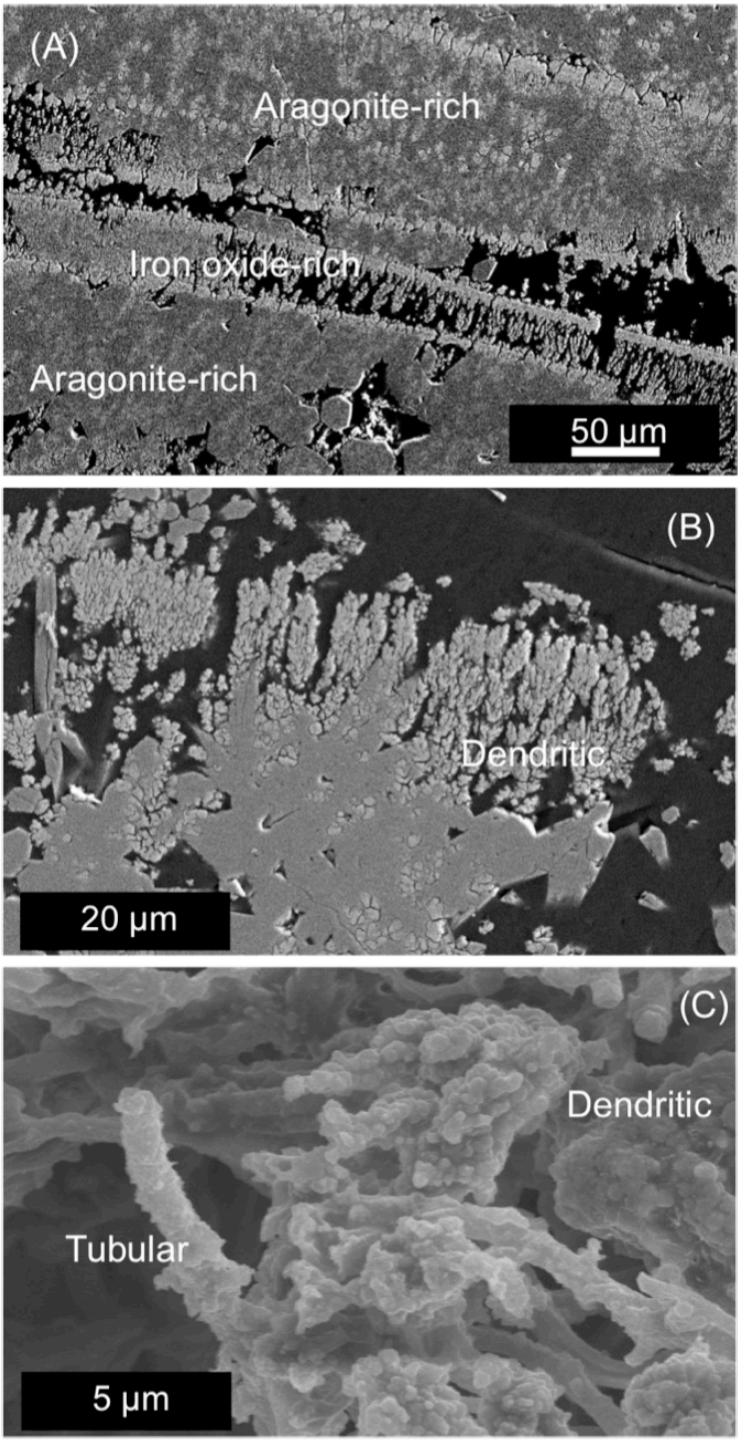


Figure 4: Electron microscopy imaging of precipitates from source pool. Back-scattered electron images collected by SEM of thin section samples of solids collected from

the travertine wall of the source pool (Figure 2D). This sample shows alternating laminations of aragonite-rich zones and iron oxide-rich zones. A) overview image illustrates alternation of carbonate- and iron oxide-rich layers. B) enlarged image of iron oxide-rich zone of A), showing aggregation of iron oxide nanoparticles into a dendritic texture. C) image of dendritic iron oxide aggregates following removal of carbonate phases via treatment with HCl, revealing sheath-like tubular and amorphous particulate iron oxide morphotypes.

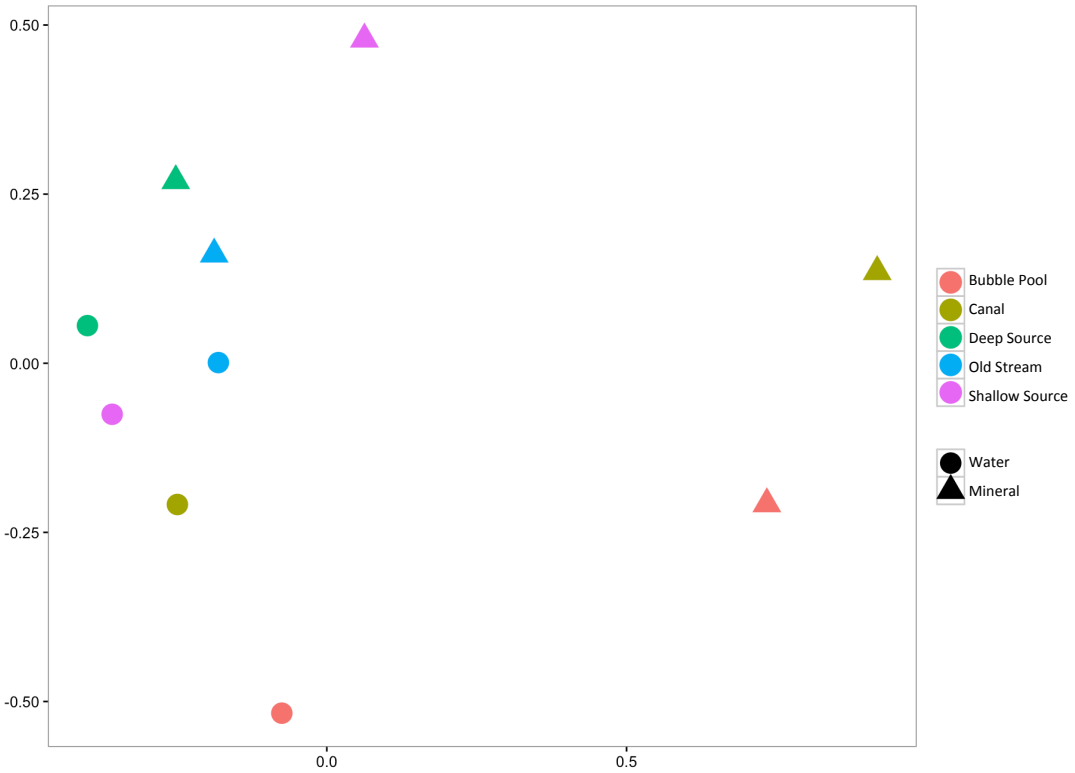


Figure 5: Multidimensional scaling analysis of OHK samples. Each point represents the recovered microbial community from a given sample, with sites identified by color and sample type by shape. Points close to each other in this two dimensional space share a similar community composition. Relative abundance data were transformed by the 4th root to down-weight the effect of abundant taxa in the samples. Water samples cluster together, and with Source pool mineral samples, while the Bubble Pool Water and the other Mineral samples stand out along different curves in this space. Stress value is 0.0465.

T	44.3°C
pH	6.8
DO	<15 μM
Fe^{2+}	114 μM
NH_3/NH_4^+	22 μM
Cl^-	38 mM
SO_4^{--}	6.5 mM
NO_3^-	b.d.
NO_2^-	b.d.
HPO_4^-	b.d.
TOC	b.d.

Table 1: Geochemical characteristics of OHK source water.

	Deep Source Mineral	Deep Source Water	Shallow Source Mineral	Shallow Source Water	Old Stream Mineral	Old Stream Water	Canal Mineral	Canal Water	Bubble Pool Mineral	Bubble Pool Water
Reads:	6456	20067	6512	25306	27454	19052	2176	17893	3608	12601
Observed OTUs (99%)	906	1928	1045	2450	1775	1842	740	2455	857	1595
Good Coverage (99%):	0.9035	0.9388	0.8948	0.941	0.9615	0.9423	0.7468	0.9076	0.8287	0.911
Shannon Index (99%):	5.9544	6.0339	7.6849	6.4785	6.8302	6.3403	8.113	6.8482	7.2747	5.896
Inverse Simpson (99%):	9.4959	12.9196	62.6337	11.5082	37.0168	11.4379	105.3153	12.7881	32.6025	8.6475
Observed OTUs (97%)	431	895	572	1059	747	869	490	1281	555	910
Goods Coverage (97%)	0.9588	0.9737	0.9459	0.9768	0.9834	0.9738	0.8359	0.9529	0.8911	0.951
Shannon Index (97%)	4.7603	4.6789	6.4739	5.0554	5.5785	5.0959	7.0603	5.4767	6.1736	4.5711
Inverse Simpson (97%):	6.5623	8.4309	33.1331	7.5288	22.6658	7.7054	55.6231	8.1763	18.5504	5.567

Table 2: Diversity metrics of OHK sequencing. Diversity metrics calculated for both 99% and 97% sequence identity cutoffs for assigning OTUs.

Taxon	Deep Source Mineral	Deep Source Water	Shallow Source Mineral	Shallow Source Water	Old Stream Mineral	Old Stream Water	Canal Mineral	Canal Water	Bubble Pool Mineral	Bubble Pool Water	Average
Bacteria;_Proteobacteria;_Betaproteobacteria;_Nitrosomonadales;_Gallionellaceae	36.94%	18.23%	1.09%	34.05%	3.27%	33.57%	0.46%	31.77%	0.44%	23.58%	18.34%
Bacteria;_Proteobacteria;_Zetaproteobacteria;_Mariprofundales;_Mariprofundaceae	0.65%	1.32%	0.05%	0.52%	1.51%	0.52%	1.01%	0.58%	0.17%	0.17%	0.65%
Bacteria;_Cyanobacteria;_Cyanobacteria;_SubsectionIII;_FamilyI	0.15%	0.24%	2.50%	4.46%	0.21%	1.41%	0.46%	2.30%	0.33%	37.45%	4.95%
Bacteria;_Cyanobacteria;_Cyanobacteria;_SubsectionV;_FamilyI	0.08%	0.11%	0.09%	5.16%	0.01%	0.31%	0.00%	1.29%	0.06%	0.20%	0.73%
Bacteria;_Proteobacteria;_Alphaproteobacteria;_Rhodospirillales;Other	0.09%	0.05%	6.68%	0.04%	2.57%	0.17%	0.09%	0.13%	0.19%	0.03%	1.00%
Bacteria;_Proteobacteria;_Betaproteobacteria;_Rhodocyclales;_Rhodocyclaceae	0.94%	0.55%	0.15%	0.42%	2.61%	3.31%	0.00%	0.99%	0.53%	0.38%	0.99%
Bacteria;_Proteobacteria;_Alphaproteobacteria;_Rhodobacterales;_Rhodobacteraceae	0.11%	0.34%	1.80%	0.47%	2.85%	0.26%	0.74%	0.87%	0.25%	0.14%	0.78%
Bacteria;_Proteobacteria;_Alphaproteobacteria;_Rhodospirillales;_Rhodospirillaceae	0.03%	0.01%	3.13%	0.04%	0.07%	0.06%	0.09%	0.30%	0.19%	0.29%	0.42%
Bacteria;_Chloroflexi;_Chloroflexia;_Chloroflexales;_Chloroflexaceae	0.00%	0.00%	0.05%	0.02%	0.00%	0.00%	1.19%	0.27%	0.83%	0.12%	0.25%
Bacteria;_Chlorobi;_Chlorobia;_Chlorobiales;_OPB56	0.09%	0.21%	0.06%	0.03%	0.17%	0.30%	0.14%	0.17%	0.03%	0.02%	0.12%
Bacteria;_Proteobacteria;_Gammaproteobacteria;_Chromatiales;_Chromatiaceae	0.42%	0.00%	0.00%	0.04%	0.00%	0.18%	0.00%	0.01%	0.03%	0.21%	0.09%
Bacteria;_Proteobacteria;_Gammaproteobacteria;_Xanthomonadales;_Xanthomonadaceae	0.22%	0.61%	3.06%	0.41%	10.01%	0.30%	9.33%	1.70%	24.89%	1.39%	5.19%
Bacteria;_Chloroflexi;_Anaerolineae;_Anaerolineales;_Anaerolineaceae	7.13%	10.75%	10.10%	3.52%	2.32%	0.91%	1.47%	3.67%	2.22%	0.60%	4.27%
Bacteria;_Proteobacteria;_Betaproteobacteria;_Burkholderiales;_Comamonadaceae	1.50%	1.88%	2.03%	2.35%	12.12%	1.51%	1.42%	4.15%	2.52%	4.38%	3.39%
Bacteria;_Nitrospirae;_Nitrospira;_Nitrospirales;_Nitrospiraceae	0.39%	1.19%	0.25%	5.07%	4.65%	12.81%	0.18%	3.82%	0.03%	2.10%	3.05%
Archaea;_Thaumarchaeota;_Marine_Group_I;_o;_f	0.99%	0.23%	0.00%	0.09%	0.01%	0.98%	0.00%	0.44%	0.00%	0.04%	0.28%
Bacteria;_Chlorobi;_Ignavibacteria;_Ignavibacteriales;Other	6.72%	34.25%	3.67%	5.38%	12.95%	2.23%	0.37%	11.85%	4.49%	1.66%	8.36%
Bacteria;_Proteobacteria;_Deltaproteobacteria;_Bdellovibrionales;_Bacteriovoracaceae	2.29%	0.97%	0.21%	4.31%	0.21%	11.38%	0.14%	4.21%	0.08%	3.75%	2.76%
Bacteria;_Candidate_division_OP11;_c;_o;_f	0.15%	1.73%	0.14%	0.46%	2.09%	0.42%	6.34%	0.25%	11.03%	0.06%	2.27%
Bacteria;_Bacteroidetes;_Sphingobacteriia;_Sphingobacteriales;_env.OPS_17	4.20%	3.28%	0.25%	7.09%	0.14%	0.68%	1.24%	3.85%	1.14%	0.64%	2.25%
Bacteria;_Proteobacteria;_Gammaproteobacteria;_Enterobacteriales;_Enterobacteriaceae	3.44%	1.01%	3.49%	1.16%	0.69%	1.47%	4.18%	0.39%	2.19%	0.53%	1.86%
Bacteria;_Proteobacteria;_Deltaproteobacteria;_Desulfuromonadales;_Geobacteraceae	0.02%	0.79%	0.00%	0.03%	0.14%	0.03%	0.05%	0.06%	0.00%	0.14%	0.13%
Bacteria;_Cyanobacteria;_ML635J-21;_o;_f	0.03%	0.06%	0.02%	0.06%	0.00%	0.34%	0.00%	0.02%	0.00%	0.23%	0.08%
Bacteria;_Cyanobacteria;_SHA-109;_o;_f	0.00%	0.00%	0.10%	0.00%	0.90%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%

Table 3: Relative abundance of taxa to the Family level. Overall 10 most abundant taxa listed, as well as other taxa of interest mentioned in the text and SI, ordered by relative abundance averaged across all samples. Taxa color coded by putative metabolism: red for aerobic iron oxidizers, purple for anoxygenic phototrophs, green for oxygenic phototrophs, blue for nitrifiers, and yellow for taxa with members performing a diverse range of metabolisms that can't be resolved with taxonomic resolution of available data.

References:

1. Bekker, A. et al. Iron Formations: Their Origins and Implications for Ancient Seawater Chemistry. Treatise on Geochemistry: Second Edition 9, (2013).
2. Beukes, N. J. "Sedimentology of the Kuruman and Griquatown iron-formations, Transvaal Supergroup, Griqualand West, South Africa."Precambrian Research 24.1 (1984): 47-84.
3. Beukes NJ and Klein C (1990) Geochemistry and sedimentology of a facies transition— from microbanded to granular iron-formation—in the early Proterozoic Transvaal Supergroup, South Africa. Precambrian Research 47: 99–139.
4. Beukes, N.J., Klein, C., Kaufman, A.J., and Hayes, J.M., 1990, Carbonate petrography, kerogen distribution, and carbon and oxygen isotope variations in and early Proterozoic transition from limestone to iron forma- tion deposition, Transvaal Supergroup, South Africa: Economic Geology and the Bulletin of the Society of Economic Geologists, v. 85, p. 663–690.
5. Bird, L. J., Bonnefoy, V. & Newman, D. K. Bioenergetic challenges of microbial iron metabolisms. Trends Microbiol. 19, 330–340 (2011).
6. Blöthe, M., and E. E. Roden. 2009. Microbial iron redox cycling in a circumneutral-pH groundwater seep. Appl. Environ. Microbiol. 75:468-473.
7. Bosak, T. et al. Cyanobacterial diversity and activity in modern conical microbialites. Geobiology 10, 384–401 (2012).

8. Bryant DA, Liu Z. 2013. Green Bacteria: Insights into Green Bacterial Evolution through Genomic Analyses. In Genome Evolution of Photosynthetic Bacteria, pp. 99–150. Academic Press
9. Byrne-Bailey, K. G. et al. Completed genome sequence of the anaerobic iron-oxidizing bacterium Acidovorax ebreus strain TPSY. J. Bacteriol. 192, 1475–1476 (2010).
10. Cairns-Smith, A.G., 1978, Precambrian solution photochemistry, inverse segregation, and banded iron formation: Nature, v. 276, p. 807–808.
11. Canfield, D. E. "A new model for Proterozoic ocean chemistry." Nature396.6710 (1998): 450-453.
12. Caporaso, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6, 1621–1624 (2012).
13. Caporaso, J. Gregory, et al. "QIIME allows analysis of high-throughput community sequencing data." Nature methods 7.5 (2010): 335-336.
14. Chan, C., Emerson, D. & Luther, G. W. The role of microaerophilic Fe-oxidizing microorganisms in producing banded iron formations. Geobiology (2016).
15. CLOUD P. 1973. Paleoecological significance of the banded iron-formation. Economic Geology 68, 1135– 43.
16. Crossey, Laura J., et al. "Continental smokers couple mantle degassing and distinctive microbiology within continents." *Earth and Planetary Science Letters* 435 (2016): 22-30.

17. Crowe, S. A. et al. Photoferrotrophs thrive in an Archean Ocean analogue. Proc. Natl. Acad. Sci. 105, 15938–15943 (2008).

18. Davenport C, Ussery DW, Tümmler B. 2010. Comparative genomics of green sulfur bacteria. Photosynth. Res. 104(2-3):137–52

19. DeLong, Edward F., et al. The Prokaryotes: Alphaproteobacteria and Betaproteobacteria. Ed. Eugene Rosenberg. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.

20. Di Rienzi SC, Sharon I, Wrighton KC, Koren O, Hug LA, et al. 2013. The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. eLife. 2:e01102

21. Drever, James I. "Geochemical model for the origin of Precambrian banded iron formations." Geological Society of America Bulletin 85.7 (1974): 1099-1106.

22. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460-2461.

23. Ehrenreich, A., and Widdel, F., 1994, Anaerobic oxidation of ferrous iron by purple bacteria, a new type of phototrophic metabolism: Applied and Environmental Microbiology, v. 60, p. 4517–4526

24. Emerson, D. & Moyer, C. Isolation and characterization of novel iron-oxidizing bacteria that grow at circumneutral pH. Appl. Environ. Microbiol. 63, 4784–4792 (1997).Emerson, D. et al. A novel lineage of proteobacteria involved in formation of marine Fe-oxidizing microbial mat communities. PLoS One 2, (2007).

25. Emerson, D., Fleming, E. J. & McBeth, J. M. Iron-oxidizing bacteria: an environmental and genomic perspective. Annu. Rev. Microbiol. 64, 561–583 (2010).

26. Emerson, D. *et al*. Comparative genomics of freshwater Fe-oxidizing bacteria : implications for physiology , ecology , and systematics. *Front. Microbiol*. **4,** 1–17 (2013).

27. Emerson, Joanne B., et al. "Metagenomic analysis of a high carbon dioxide subsurface microbial community populated by chemolithoautotrophs and bacteria and archaea from candidate phyla." *Environmental microbiology* 18.6 (2016): 1686-1703.

28. España, Javier Sánchez, Esther Santofimia Pastor, and Enrique López Pamo. "Iron terraces in acid mine drainage systems: a discussion about the organic and inorganic factors involved in their formation through observations from the Tintillo acidic river (Riotinto mine, Huelva, Spain)." *Geosphere* 3.3 (2007): 133-151.

29. Fernández-Remolar, David C., and Andrew H. Knoll. "Fossilization potential of iron-bearing minerals in acidic environments of Rio Tinto, Spain: Implications for Mars exploration." Icarus 194.1 (2008): 72-85.

30. Finneran, K. T., Johnsen, C. V. & Lovley, D. R. Rhodoferax ferrireducens sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III). Int. J. Syst. Evol. Microbiol. 53, 669–673 (2003).

31. Fischer, W. W. & Knoll, A. H. An iron shuttle for deepwater silica in late Archean and early Paleoproterozoic iron formation. Bull. Geol. Soc. Am. 121, 222–235 (2009).

32. Fischer, W. W. *et al*. SQUID-SIMS is a useful approach to uncover primary signals in the Archean sulfur cycle. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 5468–73 (2014).

33. Fischer, W., Hemp, J. & Johnson, J. E. Evolution of Oxygenic Photosynthesis. Annu. Rev. Earth Planet. Sci. 44, (2016).

34. Fleming, E. J. *et al*. Hidden in plain sight : discovery of sheath-forming ,. *FEMS Micobiol Ecol* **85,** 116–127 (2013).

35. Fortney, N. W. et al. Microbial Fe(III) oxide reduction potential in Chocolate Pots hot spring, Yellowstone National Park. Geobiology (2016).

36. Fouke, Bruce W. "Hot-spring Systems Geobiology: abiotic and biotic influences on travertine formation at Mammoth Hot Springs, Yellowstone National Park, USA." *Sedimentology* 58.1 (2011): 170-219.

37. Francois, L.M., 1986, Extensive deposition of banded iron formations was possible without photosynthesis: Nature, v. 320, p. 352–354.

38. Geological Survey of Japan, AIST (ed.). 2012. Seamless digital geological map of Japan 1: 200,000. Jul 3, 2012 version. Research Information Database DB084, Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology.

39. GOOD, I.J., 1953, The population frequencies of species and the estimation of population parameters: Biometrika, v. 40, p. 237–264.

40. Harder, E.C., 1919, Iron-Depositing Bacteria and Their Geologic Relations: U.S. Geological Survey Professional Paper 113, 89 p.

41. Hegler, Florian, et al. "Influence of seasonal and geochemical changes on the geomicrobiology of an iron carbonate mineral water spring." *Applied and environmental microbiology* 78.20 (2012): 7185-7196.

42. Heising, S., Richter, L., Ludwig, W. & Schink, B. Chlorobium ferrooxidans sp. nov., a phototrophic green sulfur bacterium that oxidizes ferrous iron in coculture with a 'Geospirillum' sp. strain. Arch. Microbiol. 172, 116–124 (1999).

43. Hemp J, Ward LM, Pace LA, Fischer WW. 2015a. Draft genome sequence of Levilinea saccharolytica KIBI-1, a member of the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01357-15.

44. Hemp J, Ward LM, Pace LA, Fischer WW. 2015b. Draft genome sequence of Ardenticatena maritima 110S, a thermophilic nitrate- and iron-reducing member of the Chloroflexi class Ardenticatenia. Genome Announc 3(6):e01347-15.

45. HILL, M.O., 1973, Diversity and evenness: a unifying notation and its consequences: Ecology, v. 54, p. 427–432.

46. Hoffman, P. F., Kaufman, A., Halverson, G. & Schrag, D. A Neoproterozoic Snowball Earth. Science 281, 1342–1346 (1998).

47. Holland, Heinrich D. "The oceans; a possible source of iron in iron-formations." Economic Geology 68.7 (1973): 1169-1172.

48. Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press

49. Iino T, Mori K, Uchino Y, Nakagawa T, Harayama S, Suzuki K-I. 2010. Ignavibacterium album gen. nov., sp. nov., a moderately thermophilic anaerobic bacterium isolated from microbial mats at a terrestrial hot spring and proposal of Ignavibacteria classis nov., for a novel lineage at the periphery of green sulfur bacteria. Int. J. Syst. Evol. Microbiol. 60(Pt 6):1376–82

50. Johnson, Clark M., et al. "Iron isotopes constrain biologic and abiologic processes in banded iron formation genesis." *Geochimica et Cosmochimica Acta* 72.1 (2008): 151-169.

51. Johnson, J. E. et al. Manganese-oxidizing photosynthesis before the rise of cyanobacteria. Proc. Natl. Acad. Sci. 110, 11238–11243 (2013a).

52. Johnson, J. E. et al. Reply to Jones and Crowe: Correcting mistaken views of sedimentary geology, Mn-oxidation rates, and molecular clocks. Proc. Natl. Acad. Sci. 110, E4119–E4120 (2013b).

53. Kadnikov, Vitaly V., et al. "Genomic analysis of Melioribacter roseus, facultatively anaerobic organotrophic bacterium representing a novel deep lineage within Bacteriodetes/Chlorobi group." *PLoS One* 8.1 (2013): e53047.

54. Kappler, A. & Newman, D. K. Formation of Fe(III)-minerals by Fe(II)-oxidizing photoautotrophic bacteria. Geochim. Cosmochim. Acta 68, 1217–1226 (2004).

55. Kappler, A., Pasquero, C., Konhauser, K. O. & Newman, D. K. Deposition of banded iron formations by anoxygenic phototrophic Fe(II)-oxidizing bacteria. Geology 33, 864–865 (2005).

56. KASAMA T. & MURAKAMI T. 2001. The effect of micro- organisms on Fe precipitation rate at neutral pH. Chemical Geology 180, 117–28.

57. Katoh, Kazutaka, et al. "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform." Nucleic acids research 30.14 (2002): 3059-3066.

58. Kirschvink, J. L. Late Proterozoic low-latitude global glaciation: the snowball Earth. In The Proterozoic Biosphere 52, 51–52 (1992).

59. Kiss, H. et al. Complete genome sequence of the filamentous gliding predatory bacterium Herpetosiphon aurantiacus type strain (114-95T). Stand. Genomic Sci. 5, 356–370 (2011).

60. Klatt, C. G. et al. Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. ISME J. 5, 1262–1278 (2011).

61. Klein, C. Some Precambrian banded iron-formations (BIFs) from around the world: Their age, geologic setting, mineralogy, metamorphism, geochemistry, and origin. Am. Mineral. 90, 1473–1499 (2005).

62. Klueglein, N., and A. Kappler. "Abiotic oxidation of Fe (II) by reactive nitrogen species in cultures of the nitrate-reducing Fe (II) oxidizer Acidovorax sp.

BoFeN1–questioning the existence of enzymatic Fe (II) oxidation."Geobiology 11.2 (2013): 180-190.

63. Konhauser, K. O. et al. Could bacteria have formed the Precambrian banded iron formations? Geology 30, 1079–1082 (2002).
64. Konhauser, K. O., Newman, D. K. & Kappler, A. The potential significance of microbial Fe (III) reduction during deposition of Precambrian banded iron formations. Geobiology 3, 167–177 (2005).
65. Kopf, S. H., Henny, C. & Newman, D. K. Ligand-enhanced abiotic iron oxidation and the effects of chemical versus biological iron cycling in anoxic environments. Environ. Sci. Technol. 47, 2602–2611 (2013).
66. Kucera, S. & Wolfe, R. S. A Selective Enrichment Method for Gallionella Ferruginea. J. Bacteriol. 74, 344–349 (1957).
67. Ley, Ruth E., et al. "Obesity alters gut microbial ecology." Proceedings of the National Academy of Sciences of the United States of America 102.31 (2005): 11070-11075.
68. Li, Yi-Liang, et al. "Experimental low-grade alteration of biogenic magnetite indicates microbial involvement in generation of banded iron formations." *Earth and Planetary Science Letters* 361 (2013): 229-237.
69. Liu Z, Frigaard N-U, Vogl K, Iino T, Ohkuma M, et al. 2012a. Complete Genome of Ignavibacterium album, a Metabolically Versatile, Flagellated, Facultative Anaerobe from the Phylum Chlorobi. Front. Microbiol. 3:185

70. Liu Z, Klatt CG, Ludwig M, Rusch DB, Jensen SI, et al. 2012b. “Candidatus Thermochlorobacter aerophilum:” an aerobic chlorophotoheterotrophic member of the phylum Chlorobi defined by metagenomics and metatranscriptomics. ISME J. 6(10):1869–82

71. Mori, J. F. *et al*. Iron encrustations on filamentous algae colonized by Gallionella-related bacteria in a metal-polluted freshwater stream. *Biogeosciences* **12,** 5277–5289 (2015).

72. Morris, R. C. "A textural and mineralogical study of the relationship of iron ore to banded iron-formation in the Hamersley iron province of Western Australia." Economic Geology 75.2 (1980): 184-209.

73. Neubauer, S. C. et al. Life at the energetic edge: kinetics of circumneutral Fe oxidation by lithotrophic iron oxidizing bacteria isolated from the wetland plant rhizosphere. Appl. Environ. Microbiol. 68, 3988–3995 (2002).

74. Nunoura, Takuro, et al. "Isolation and characterization of a thermophilic, obligately anaerobic and heterotrophic marine Chloroflexi bacterium from a Chloroflexi-dominated microbial community associated with a Japanese shallow hydrothermal system, and proposal for Thermomarinilinea lacunofontalis gen. nov., sp. nov." *Microbes and Environments* 28.2 (2013): 228-235.

75. Oksanen, Jari, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens and Helene Wagner (2016). vegan: Community Ecology Package. R package version 2.3-5. https://CRAN.R-project.org/package=vegan

76. Overmann J. 2008. Green nonsulfur bacteria. In Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Chichester, United Kingdom.

77. Overmann, Jörg, and Ferran Garcia-Pichel. "The phototrophic way of life." The Prokaryotes. Springer Berlin Heidelberg, 2013. 203-257.

78. Pace LA, Hemp J, Ward LM, Fischer WW. 2015. Draft genome of Thermanaerothrix daxensis GNS-1, a thermophilic facultative anaerobe from the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01354-15.

79. PARADA, A., NEEDHAM, D.M., AND FUHRMAN, J.A., 2015, Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time-series and global field samples: Environmental Microbiology, v. 18, p. 1403–1414.

80. PARENTEAU, M. N. & CADY, S. L. Microbial Biosignatures in Iron-Mineralized Phototrophic Mats At Chocolate Pots Hot Springs, Yellowstone National Park, United States. *Palaios* **25,** 97–111 (2010).

81. Pierson, B. K., Parenteau, M. N. & Griffin, B. M. Phototrophs in high-iron-concentration microbial mats: Physiological ecology of phototrophs in an iron-depositing hot spring. Appl. Environ. Microbiol. 65, 5474–5483 (1999).

82. Pierson, B. K. & Parenteau, M. N. Phototrophs in high iron microbial mats: Microstructure of mats in iron-depositing hot springs. FEMS Microbiol. Ecol. 32, 181–196 (2000).

83. Podosokorskaya OA, Kadnikov VV, Gavrilov SN, Mardanov AV, Merkel AY, et al. 2012. Characterization of Melioribacter roseus gen. nov., sp. nov., a novel

facultatively anaerobic thermophilic cellulolytic bacterium from the class Ignavibacteria, and a proposal of a novel bacterial phylum Ignavibacteriae. Environ. Microbiol.

84. Podosokorskaya, Olga A., et al. "Ornatilinea apprima gen. nov., sp. nov., a cellulolytic representative of the class Anaerolineae." *International journal of systematic and evolutionary microbiology* 63.1 (2013): 86-92.
85. Posth, N. R., Konhauser, K. O. & Kappler, A. Microbiological processes in banded iron formation deposition. Sedimentology 60, 1733–1754 (2013).
86. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments. Plos One 5(3).
87. Quast, Christian, et al. "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools." Nucleic acids research41.D1 (2013): D590-D596.
88. R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
89. Rantamäki, S. et al. Oxygen produced by cyanobacteria in simulated Archaean conditions partly oxidizes ferrous iron but mostly escapes—conclusions about early evolution. Photosynth. Res. 1–9 (2016).
90. Rasmussen, B., Meier, D. B., Krapež, B. & Muhling, J. R. Iron silicate microgranules as precursor sediments to 2.5-billion-yearold banded iron formations. Geology 41, 435–438 (2013).

91. Rasmussen, B., Krapež, B., Muhling, J. R. & Suvorova, A. Precipitation of iron silicate nanoparticles in early Precambrian oceans marks Earth's first iron age. Geology 43, 303–306 (2015).

92. Rasmussen, B., Muhling, J. R., Suvorova, A. & Krapež, B. Dust to dust: Evidence for the formation of 'primary' hematite dust in banded iron formations via oxidation of iron silicate nanoparticles. Precambrian Res. 284, 49–63 (2016).

93. Roden, E., J. M. McBeth, M. Blothe, E. M. Percak-Dennett, E. J. Fleming, R. R. Holyoke et al. 2012. The microbial ferrous wheel in a neutral pH groundwater seep. Front. Microbiol. 3:172.

94. Roeselers, G. et al. Diversity of phototrophic bacteria in microbial mats from Arctic hot springs (Greenland). Environ. Microbiol. 9, 26–38 (2007).

95. Sekiguchi, Y. et al. Anaerolinea thermophila gen. nov., sp. nov. and Caldilinea aerophila gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain Bacteria at the subphylum level. Int. J. Syst. Evol. Microbiol. 53, 1843–1851 (2003).

96. Shannon, C. E. (1948) A mathematical theory of communication. The Bell System Technical Journal, 27, 379–423 and 623–656.

97. Shih, P., Hemp, J., Ward, L., Matzke, N. & Fischer, W. 2016. Crown group oxyphotobacteria postdate the rise of oxygen. Geobiology.

98. SHIMAZU M., YAMADA K., NARITA E. & IGARASHI T. 1965. Geology of the Ainai-Kosaka-Oyu areas, Akita Prefecture. Bulletin of the Geological Survey of Japan 16, 22–30

99. SIMPSON, E.H., 1949, Measurement of diversity: Nature, v. 163, p. 688.

100. Singer, P. C. & Stumm, W. Acid Mine Drainage: The Rate-Determining Step. *Science (80-.).* **167,** 1121–1123 (1970).

101. Soo RM, Skennerton CT, Sekiguchi Y, Imelfort M, Paech SJ, et al. 2014. An expanded genomic representation of the phylum cyanobacteria. Genome Biol. Evol. 6(5):1031–45

102. Soo RM, Woodcroft BJ, Parks DH, Tyson GW, Hugenholtz P. 2015. Back from the dead; the curious tale of the predatory cyanobacterium Vampirovibrio chlorellavorus. PeerJ. 3:e968

103. Soo, Rochelle M., et al. "On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria." *Science* 355.6332 (2017): 1436-1440.

104. Stamps BW, Corsetti FA, Spear JR, Stevenson BS. 2014. Draft Genome of a Novel Chlorobi Member Assembled by Tetranucleotide Binning of a Hot Spring Metagenome. Genome Announc. 2(5):e00897–14 – e00897–14

105. Stookey, Lawrence L. "Ferrozine---a new spectrophotometric reagent for iron." Analytical chemistry 42.7 (1970): 779-781.

106. Stumm, W. & Lee, F. G. Oxygenation of ferrous iron. *Ind. Eng. Chem.* **53,** 143–146 (1961).

107. Swanner, E. D. et al. Modulation of oxygen production in Archaean oceans by episodes of Fe(II) toxicity. Nat. Geosci. 8, 126–130 (2015).

108. Takashima, C., Okumura, T., Nishida, S., Koike, H. & Kano, A. Bacterial symbiosis forming laminated iron-rich deposits in Okuoku-hachikurou hot spring, Akita Prefecture, Japan. Isl. Arc 20, 294–304 (2011).

109. Tosca, Nicholas J., Stephen Guggenheim, and Peir K. Pufahl. "An authigenic origin for Precambrian greenalite: Implications for iron formation and the chemistry of ancient seawater." Geological Society of America Bulletin 128.3-4 (2016): 511-530.

110. Trembath-Reichert E, Ward LM, Slotznick SP, Bachtel SL, Kerans C, Grotzinger JP, Fischer WW (2016) Gene sequencing microbial community analysis of mat morphologies, Caicos Platform, British West Indies, Journal of Sedimentary Research, in press.

111. Trouwborst, R. E., Johnston, A., Koch, G., Luther, G. W. & Pierson, B. K. Biogeochemistry of Fe(II) oxidation in a photosynthetic microbial mat: Implications for Precambrian Fe(II) oxidation. Geochim. Cosmochim. Acta 71, 4629–4643 (2007).

112. van Veen, W. L., Mulder, E. G. & Deinema, M. H. The Sphaerotilus-Leptothrix group of bacteria. Microbiol. Rev. 42, 329–356 (1978).

113. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microb 73(16): 5261-5267.

114. Ward, L. M., Kirschvink, J. L. & Fischer, W. W. 2016. Timescales of Oxygenation Following the Evolution of Oxygenic Photosynthesis. Orig. Life Evol. Biosph. 46(1) pp51-65.

115. Ward LM, Hemp J, Pace LA, Fischer WW. 2015a. Draft genome sequence of Leptolinea tardivitalis YMTK-2, a mesophilic anaerobe from the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01356-15.

116. Ward LM, Hemp J, Pace LA, Fischer WW. 2015b. Draft genome sequence of Herpetosiphon geysericola GC-42, a nonphototrophic member of the Chloroflexi class Chloroflexia. Genome Announc 3(6):e01352-15

117. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009.

118. Widdel, F., Schnell, S., Heising, S., Ehrenreich, A., Assmus, B., and Schink, B., 1993, Ferrous iron oxidation by anoxygenic phototrophic bacteria: Nature, v. 362, p. 834–836.

119. Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H,Kamagata Y. 2006. Anaerolinea thermolimosa sp. nov., Levilinea saccharolyticagen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov.and Caldilineae classis nov. in the bacterial phylum Chloroflexi. Int J Syst Evol Microbiol 56:1331–1340.

120. Yamada, T. & Sekiguchi, Y. Cultivation of uncultured chloroflexi subphyla: significance and ecophysiology of formerly uncultured chloroflexi ‘subphylum i’

with natural and biotechnological relevance. Microbes Environ. 24, 205–216 (2009).

Supplemental Information:

Materials and methods:

Rarefaction plots were produced using QIIME version 1.8.0 (Caporaso et al., 2010) (Supplemental Figure 1). We used the UniFrac distance metric (Lozupone and Knight 2005) to assess the microbial community phylogenetic similarity of microbial communities (Supplemental Table 1).

Light microscopy images (Supplementary Figure 2) were taken on a Zeiss Axio Imagers 2 Upright microscope (Zeiss, Germany) with 40x and 100x objective lens. Epi-fluorescent images of DAPI (4'6-diamidino-2-phenylindole)-stained samples were taken with an excitation wavelength below 395 nm and an emission wavelength between 420 and 470 nm. The autofluorescence of Cyanobacteria was detected by exposing the sample to a wavelength of between 395 and 440 nm and detecting the emission at a wavelength of 470 nm. Images of DAPI fluorescence and autofluorescence were overlain using the FIJI software package (http://pacific.mpi-cbg.de).

Secular and paleoenvironmental changes in the redox state of iron in iron formations:

BIF deposits are compositionally, texturally, and sedimentologically diverse and consequently may not have a single causal mechanism of origin (Fischer and Knoll 2009). Broadly, BIFs can be grouped into three categories based on age of deposition, each with characteristic mineralogical and redox compositions and hypothesized formation mechanisms: Archean (and early Paleoproterozoic) BIF, which formed before the GOE, middle to late Paleoproterozoic BIF, which formed after the GOE and before ~1.8 Ga, and

Neoproterozoic BIF, which formed later in time, principally ~0.8-0.6 Ga, associated with Cryogenian glacial intervals (Hoffman et al. 1998, Kirschvink 1992). Archean BIFs consist primarily of well-laminated, sediment-starved and often deep water facies, consistent with deposition in an anoxic environment where deep waters were enriched in Fe(II) and could be oxidized either biotically or abiotically during mixing into the photic zone. This style of BIF exists throughout Archean and early Paleoproterozoic successions until the GOE. The iron minerals that comprise this category of BIF are made up largely of iron carbonates and silicates (Beukes 1984, Beukes and Klein 1990, Bekker et al. 2013, Rasmussen et al. 2013, Rasmussen et al. 2015, Rasmussen et al. 2016). These BIFs contain a high proportion of reduced iron, reflecting some combination of primary deposition of ferrous minerals (Rasmussen et al. 2013, Rasmussen et al. 2015, Rasmussen et al. 2016) and deposition of iron oxides followed by early microbial re-reduction associated with respiration of organic carbon (Konhauser et al. 2005, Fischer and Knoll 2009, Fischer et al. 2009, Bontognali et al. 2013). The most abundant iron oxide phase in Archean BIFs is magnetite, which has a late diagenetic origin where it occurs (Han 1978, Beukes 1984, Beukes et al. 1990, Fischer and Knoll 2009). Paleoproterozoic BIFs range from ~2.3 Ga to 1.8 Ga, and commonly consist of more granular and oolitic facies with current-generated structures such as dune- and ripple-scale cross-stratification, reflecting deposition in relatively higher energy, shallower water environments. This is consistent with oxidation via atmospheric O_2 of iron sourced from anoxic basinal waters and pore fluids. Paleoproterozoic BIFs typically contain higher abundances of hematite, magnetite, and other ferric iron minerals than observed in Archean BIFs (Klein 2005, Fischer and Knoll 2009, Bekker et al. 2013),

consistent with deposition in a more oxidizing world where carbon remineralization is driven more by aerobic respiration than by iron reduction. Furthermore, Paleoproterozoic BIFs contain much lower organic carbon than Archean BIFs (Fischer et al. 2009), suggesting that organic matter remineralization was much more efficient, as would be expected for an aerobic respiration-driven system. BIFs largely disappear from middle Proterozoic basins, likely driven by the increased abundances of both oxygen and sulfur in the fluid Earth (Canfield 1998), until they reappear between ~750 and 635 Ma in association with Cryogenian glacial intervals (Kirschvink et al. 1992, Hoffman et al. 1998). Neoproterozoic BIFs are highly unique in that they are made up almost entirely by ferric iron minerals (primarily hematite) (Klein 2005).

Additional discussion of the microbiology at OHK:

Interestingly, the water and minerals at OHK contain a high proportion of Xanthomonadaceae. Members of the Xanthamonadaceae family within the Gammaproteobacteria are typically plant-associated bacteria, particularly pathogens (Starr 1981). The Xanthomonadaceae sequences present at OHK are most similar (95% identity) to *Thermomonas*, a genus that includes members such as *Thermomonas hydrothermalis* isolated from hot springs where they act as aerobic heterotrophs (Alves et al. 2003).

While concentrations of fixed nitrogen species at OHK were low (no detectable nitrate or nitrite, and a maximum of 0.37 mg/l ammonia in the source pool)(Table 1), sequences associated with diverse nitrogen cycling organisms were recovered (Table 3). These include members of the Nitrospiraceae at up to 12.8% in the Old Stream Mineral

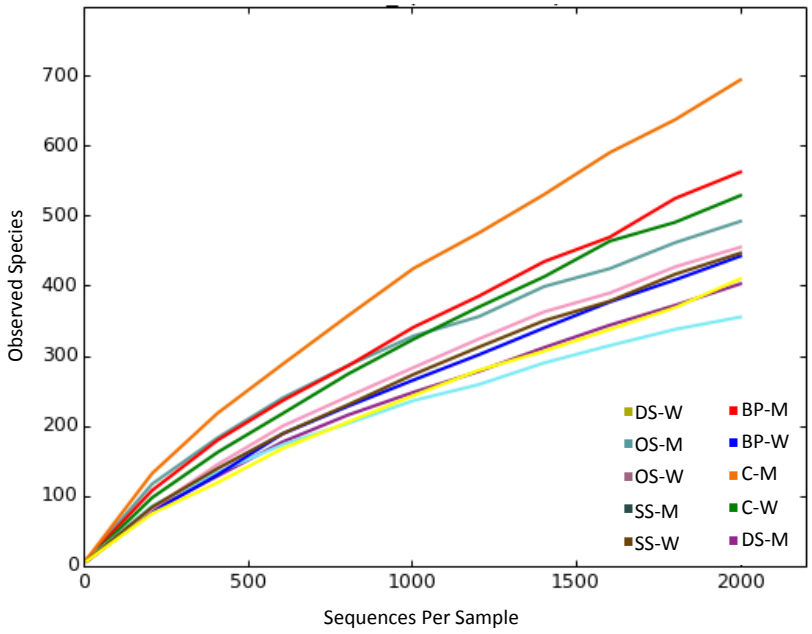
sample. The Nitrospiraceae, including the genus *Nitrospira*, are typically characterized as nitrite oxidizers (Bock and Wagner 2006) but are also one of the first groups known to possess the capacity for complete nitrification (comammox) (Daims et al. 2015, van Kessel et al. 2015).

Members of the Marine Group 1 Thaumarchaeota were the most abundant archaeal taxon observed at OHK (up to ~1% in the Deep Source Mineral and Old Source Water samples)(Table 3). This group is commonly characterized by aerobic ammonia oxidation in open ocean environments (e.g. Spang et al. 2010, Swan et al. 2014). The ammonia-oxidizing archaea are typically adapted to much lower ammonia concentrations than ammonia oxidizing bacteria (Martens-Habbena et al. 2009, Hatzenpichler 2012), consistent with their presence at OHK where concentrations of fixed nitrogen are low.

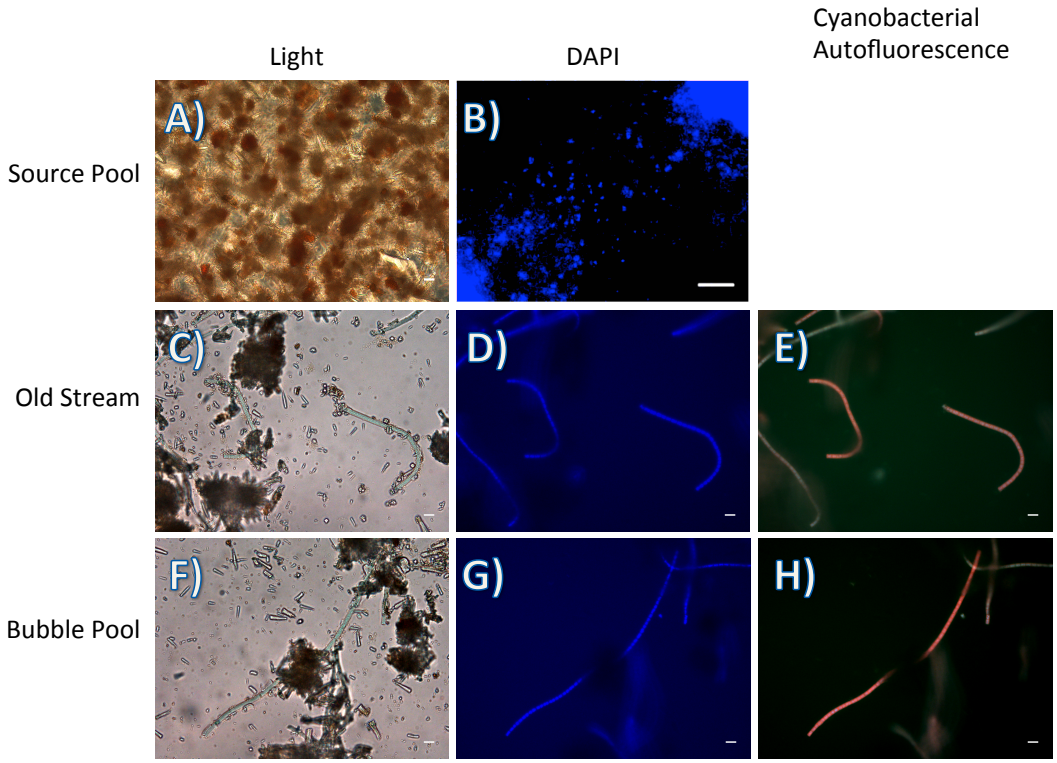
Although we observed nitrifying organisms at OHK, their metabolic byproducts—nitrite and nitrate—were not detected (Table 1). This may reflect the rapid reaction of oxidized nitrogen species with ferrous iron, which can proceed either via biological or abiotic mechanisms (Straub et al. 1996, Klueglein and Kappler 2013, Kopf et al. 2013). Like with many cycles that contain reactive intermediates, the low standing stocks of fixed nitrogen species may not preclude its overall importance to biogeochemical cycling, and thus a component of iron oxidation at OHK may occur through a cryptic nitrogen cycle. Even if this is the case, the ultimate source of oxidizing potential would nonetheless originate in atmospheric O_2 used for nitrification. However, it is interesting to note that many known iron oxidizers are nested within nitrifying clades (e.g. *Gallionella* within Nitrosomonadales, *Leptospirillum* within Nitrospirae)—hinting at metabolic, ecological,

and evolutionary links between these chemoautotrophic metabolisms. This also leaves open the possibility that the Nitrospiraceae sequences observed here are associated not with nitrite oxidizers as is typically seen, but instead could represent a group of iron oxidizers within this clade.

Members of the Candidate Phylum OP11 are abundant at OHK, particularly in the Bubble Pool mineral fraction where they made up ~11% of sequence reads (Table 3). OP11, also known as Microgenomates, are a clade within the enigmatic Candidate Phyla Radiation of bacteria (Brown et al. 2015, Eloe-Fadrash et al. 2016, Hug et al. 2016). Though members of OP11 thus far remain uncultured, assembly and binning of metagenomic sequence data suggests that these organisms are anaerobes with a widespread capacity for fermentation-based metabolism (Wrighton et al. 2012).



Supplemental Figure 1: Subsampling (rarefaction) diversity curves of OHK samples. Sampled to a depth of 2000 reads per sample. The microbial diversity at OHK is fairly well sampled, even at this low level of sampling depth, particularly near the source where diversity is lower overall. Downstream samples, particularly the Canal Mineral sample, are more diverse and additional sequencing depth would likely reveal additional diversity.



Supplemental Figure 2: Light and epifluorescence microscopy images of representative OHK hot spring samples. Top row (A&B): Source pool; middle row (C-E): Old Stream; bottom row (F-H): Bubble Pool; Left column: light microscopy showing mineral grains and associated microbial cells; middle column: fluorescence image of same samples, showing DAPI-stained cells in blue; right column: same samples showing cyanobacterial autofluorescence. Cyanobacterial filaments are apparent in Old Stream and Bubble Pool samples, but do not appear in Source Pool samples, consistent with molecular data. Scale bars are 10 μm.

	OS-W	C-W	SS-W	DS-M	BP-W	SS-M	DS-W	OS-M	BP-M	C-M
OS-W	0	0.2719	0.3201	0.2891	0.5340	0.5710	0.4676	0.3811	0.5826	0.6008
C-W		0	0.2197	0.2568	0.5317	0.5442	0.3974	0.3354	0.5685	0.6066
SS-W			0	0.3088	0.5247	0.6120	0.4710	0.4174	0.6473	0.6919
DS-M				0	0.5776	0.5463	0.3378	0.3659	0.5401	0.5582
BP-W					0	0.8317	0.7398	0.6608	0.8481	0.8715
SS-M						0	0.5862	0.5706	0.5464	0.5053
DS-W							0	0.4050	0.5578	0.5773
OS-M								0	0.5156	0.5568
BP-M									0	0.3697
C-M										0

Supplemental Table 1:

Weighted Unifrac matrix, showing the dissimilarity between samples, incorporating both differences in presence/absence as well as relative abundance of taxa (color coded to highlight relative dissimilarity—dark green is <0.33 dissimilarity, pale green >0.33 and <0.5, yellow >0.5 and <0.66, orange >0.66). Samples from the same location tend to be relatively similar, and water samples tend to be relatively similar. The Bubble Pool Water sample was most dissimilar to the other samples. OS, Old Stream; C, Canal; SS, Shallow Source Pool; DS, Deep Source Pool; BP, Bubble Pool; -W, filtered water sample; -M mineral precipitate sample.

Supplemental references:

1. Alves, Marta P., et al. "Thermomonas hydrothermalis sp. nov., a new slightly thermophilic γ-proteobacterium isolated from a hot spring in central Portugal." *Systematic and applied microbiology* 26.1 (2003): 70-75.
2. Bock, Eberhard, and Michael Wagner. "Oxidation of inorganic nitrogen compounds as an energy source." The prokaryotes. Springer New York, 2006. 457-495.

3. Bontognali, T. R. R., Fischer, W. W. & Föllmi, K. B. Siliciclastic associated banded iron formation from the 3.2Ga Moodies Group, Barberton Greenstone Belt, South Africa. Precambrian Res. 226, 116–124 (2013).
4. Brown, C. T. et al. Unusual biology across a group comprisingmore than 15% of domain Bacteria Christopher. Nature 523, 208–211 (2015).
5. Daims, Holger, et al. "Complete nitrification by Nitrospira bacteria." Nature528.7583 (2015): 504-509.
6. Eloe-Fadrosh, E. A., Ivanova, N. N., Woyke, T. & Kyrpides, N. C. Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. Nat. Microbiol. 15032 (2016).
7. Fischer, W. W. et al. Isotopic constraints on the Late Archean carbon cycle from the Transvaal Supergroup along the western margin of the Kaapvaal Craton, South Africa. Precambrian Res. 169, 15–27 (2009).
8. Han, T.-M., 1978, Microstructures of magnetite as guides to its origin in some Precambrian iron-formations: Fortschritte der Mineralogie, v. 56, p. 105–142.
9. Hatzenpichler, R. Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* **78,** 7501–7510 (2012).
10. Hug, L. A. et al. A new view of the tree and life's diversity. Nat. Microbiol. (2016).
11. van Kessel, M. A. H. J. *et al*. Complete nitrification by a single microorganism. *Nature* **528,** 1–17 (2015).

12. LOZUPONE, C., AND KNIGHT, R., 2005, UniFrac: a new phylogenetic method for comparing microbial communities: Applied and Environmental Microbiology, v. 71, p. 8228–8235.

13. Martens-Habbena, W. et al. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. Nature 461, 976–979 (2009).

14. Spang, A. et al. Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. Trends Microbiol. 18, 331–340 (2010).

15. Starr, Mortimer P. "The genus Xanthomonas." The prokaryotes. Springer Berlin Heidelberg, 1981. 742-763.

16. Straub, Kristina Lotte, et al. "Anaerobic, nitrate-dependent microbial oxidation of ferrous iron." Applied and Environmental Microbiology 62.4 (1996): 1458-1460.

17. Swan, B. K. et al. Genomic and metabolic diversity of marine group i thaumarchaeota in the mesopelagic of two subtropical gyres. PLoS One 9, (2014).

18. Wrighton, K. et al. Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. Science (80-.). 337, 1661–1666 (2012).

Chapter 5

FOLLOW THE OXYGEN: COMPARATIVE HISTORIES OF PLANETARY OXYGENATION AND OPPORTUNITIES FOR LIFE

Ward, Lewis M., Vlada Stamenković, and Woodward W. Fischer. In preparation.

Abstract

Aerobic respiration—the ability to couple the reduction of molecular oxygen (O_2) to the oxidation of reduced compounds while conserving energy to drive cellular processes—is the most widespread and bioenergetically favorable metabolism on Earth today. Aerobic respiration is essential for the development of complex life, and so the presence of abundant molecular oxygen is a crucial component of planetary habitability. O_2 on Earth is supplied by oxygenic photosynthesis, but it is becoming more widely understood that other, abiotic processes may supply small amounts of O_2 on other worlds. It is not yet clear however whether these sources are sufficient for aerobic organisms. Observations of the modern Martian atmosphere and analysis of the Martian rock record suggest a history of relatively high O_2 as a result of photochemical processes, potentially overlapping with the range of oxygen concentrations used by biology. Europa, meanwhile, may have accumulated high oxygen concentrations in its subsurface ocean due to the radiolysis of ices at its surface. Recent modeling efforts have suggested that coexisting water and oxygen may be common on exoplanets, with confirmation from measurements of exoplanet atmospheres expected in the near future. In all of these cases, O_2 is thought to have accumulated through abiotic processes, independent of photosynthesis. If correct, this would have far-reaching implications for interpretations of the habitability of these

worlds, potentially allowing the development of highly-energetic aerobic respiration, and even the complex multicellular life which depends on it, without first evolving oxygenic photosynthesis, a biochemically-complex metabolism which has evolved only once in Earth history. This suggests a new approach to investigating planetary habitability: "Follow the Oxygen", in which environments with opportunities for energetic metabolisms such as aerobic respiration are targeted for investigation and life detection.

Introduction

Aerobic respiration—the ability to couple the reduction of O_2 to the oxidation of reduced compounds such as organic carbon, ferrous iron, reduced sulfur compounds, or molecular hydrogen while conserving energy to drive cellular processes—is the most widespread and bioenergetically favorable metabolism on Earth today. Aerobic respiration is also viewed as essential for the development of complex multicellular life (Catling et al. 2005), and so the presence of abundant molecular oxygen is a crucial component of planetary habitability (Cockell et al. 2016). O_2 on Earth is almost entirely supplied by oxygenic photosynthesis—a sophisticated metabolism involving the photochemical oxidation of water (Fischer et al. 2016)—but it is becoming more widely appreciated that other, abiotic processes may supply small amounts of O_2 and related compounds such as H_2O_2 (e.g. Borda et al. 2001, Liang et al. 2006, Lu et al. 2014). While it appears that these abiotic oxidant fluxes are insufficient to supply biologically useful oxygen concentrations on the Earth (Pavlov et al. 2001), differing histories and conditions on other planets could lead to sufficient accumulation of oxygen to

concentrations that would be relevant to biology. Observations of such abiotically produced oxygen oxygenation have been made on Mars, and are predicted for Europa and certain classes of exoplanets, potentially to be confirmed by upcoming observations.

Classically, the Pasteur Point—the O_2 threshold at which brewer's yeast transitions from respiration to fermentation (~3 μM)—has long defined the assumed lower oxygen limit for the viability of aerobic respiration. However, recent experiments involving diverse organisms under extremely low oxygen concentrations have demonstrated aerobic respiration and growth at far lower oxygen concentrations, expanding the viable range for aerobic life downward. Laboratory experiments with marine sponges have demonstrated that low-micromolar oxygen concentrations are sufficient to support the metabolisms of these animals (Mills et al. 2014), while similar experiments with *Escherichia coli* demonstrate growth at oxygen levels three orders of magnitude lower (Stolper et al. 2010). Meanwhile aerobic organisms including invertebrates (Sperling et al. 2013b) and microbes (e.g. Morris and Schmidt 2013) have been observed in environments that have previously been considered anoxic, indicating that this micro-aerobic respiration is ecologically meaningful in the natural environment. This suggests that more environments could be suitable for aerobic life than previously recognized. Discussion of the expanded range of environments habitable by aerobes has primarily been discussed in the context of Earth history (e.g., Zhang et al. 2016, Fischer 2016) and low oxygen environments on the modern Earth (e.g. Ulloa et al. 2012), but it also has broad relevance to a range of planetary environments, including those on Mars, Europa, and elsewhere.

O_2 is a major component of Mars' atmosphere today, where it is the fourth most abundant gas (Mahaffy et al. 2013). Though the absolute abundance of O_2 is relatively low on modern Mars due to the thin atmosphere, recent discoveries of manganese enrichments in Martian sedimentary rocks imply the existence of relatively high oxygen levels in aqueous environments sometime in Mars' past (Lanza et al 2014, Arvidson et al. 2016). These data, together with other studies (e.g. Halevy et al. 2011, Farquhar et al. 1998, Shaheen et al. 2015, and Hurowitz et al. 2010) provide support for the hypothesis that O_2 was once more abundant in Mars' atmosphere. Combined with improved understanding of the lower oxygen limits for aerobic respiration, these observations suggest that ancient Mars may have possessed sufficient oxygen to support aerobic organisms. If correct, this would have far-reaching implications for interpretations of the habitability of Mars and other worlds.

Europa is uniquely positioned in the outer solar system as a promising candidate for providing an environment habitable to aerobic life. Europa is thought to possess a liquid water ocean ~100 km deep, which is potentially oxygenated due to interactions with its icy shell, which accumulates O_2 from radiolysis of water ice at the surface (Hand et al. 2007, Hand and Brown 2013). Europa therefore provides an intriguing case study for habitability, with the interesting caveat that any Europan biosphere would be supported not by photosynthesis but instead driven by redox chemistry between oxidants produced at the surface and reductants sourced water-rock interactions at the base of the ocean (Chyba and Hand 2001). Metabolisms potentially supported by Europa include aerobic iron, sulfide, hydrogen, and methane oxidation. However, the rates of

productivity and total biomass that these processes could support are not well constrained (e.g. Ward et al. 2017, Neubauer et al. 2002), making predictions of the extent of any Europan biosphere challenging.

In this paper, we contextualize the possibility of aerobic environments on other worlds within the framework of what is known about the history of oxygen on Earth and the role of oxygen in biology. The history and modern distribution of oxygen on Mars and Europa are used as test cases to begin examining the opportunities for aerobic respiration on other planets where the history of O_2 is not contingent on the evolution of oxygenic photosynthesis. This provides a basis for considering the more general case of exoplanets, which under a range of circumstances may support oxygenated atmospheres that may be detectable with upcoming investigations. The potential for substantial abiotically-sourced oxygen in various planetary environments expands the range of potentially habitable environments, provides the opportunity for highly energetic metabolisms, and may improve the likelihood of evolution of complex life by removing evolutionary bottlenecks contingent on challenging biochemical innovations related to oxygenic photosynthesis. The impact of oxidizing environments on the origin of life is unknown, however, and could introduce challenges to the origin of life where abiotic O_2 is abundant.

Oxygen on Earth

The history of molecular oxygen on Earth is closely tied to the history of photosynthesis. Atmospheric oxygen concentrations are constrained to well below $10^{-5.7}$ atm before the Great Oxygenation Event (GOE) about 2.35 Gya, as recorded by a range

of proxies including the behavior of iron in paleosols and red beds (Holland 1984, Rye and Holland 1998), redox-sensitive detrital grains (Johnson et al. 2014), mass-independent fractionation of sulfur isotopes (Farquhar et al. 2000, Bekker et al. 2004, Johnson et al. 2013), and manganese deposits (Johnson et al. 2013, Maynard 2010). During the GOE, O_2 accumulated in the atmosphere to 1-10% present atmospheric levels (PAL) (~10^{-2}-10^{-3} atm), and has remained a significant component of the fluid Earth ever since (Lyons et al. 2014). The oxygenation of Earth's atmosphere was caused by oxygenic photosynthesis by Cyanobacteria (e.g., Falkowski 2011, Shih 2015), and may have occurred rapidly following the evolution of this metabolism (Ward et al. 2016, Shih et al. 2016, Fischer et al. 2016). The history of molecular oxygen in Earth's atmosphere is therefore tightly coupled to the history of oxygenic photosynthesis by Cyanobacteria, making it an almost entirely biological history. The availability of O_2 ultimately paved the way for the evolution of aerobic respiration (Hemp et al. 2012), endosymbiosis and radiation of the eukaryotes (Katz 2012), and complex multicellularity (Catling et al. 2005, Knoll 2011).

From the perspective of thermodynamics, aerobic respiration is by far the most energetically favorable metabolism over a tremendous range in O_2 concentrations when compared to common anaerobic respiration processes (Figure 2). For example, it typically yields 4 times more energy per electron than sulfate reduction across more than six orders of magnitude in O_2 concentrations. While thermodynamics predicts that aerobic respiration should be viable down to vanishingly low concentrations of O_2, it has remained contentious whether life could make use of these lowest concentrations, or if

kinetic constraints on the biochemistry of aerobic respiration set a threshold below which respiration is not physiologically viable. Historically, the Pasteur Point has been considered to mark the minimum oxygen concentration for aerobic respiration, below which organisms transition to fermentation (or other anaerobic metabolisms) to maintain energetic balance. The Pasteur Point corresponds to an oxygen concentration of about 1% of PAL dissolved in seawater, or about 3 μM. This is also thought to be the minimum oxygen concentration for aerobic respiration (e.g. Berkner and Marshall 1965), but a recent study demonstrated aerobic respiration by *Escherischia coli* at vastly lower oxygen concentrations, closer to 0.001% PAL (Stolper et al 2010). This greatly extends the potential range of oxygen concentrations under which aerobic respiration is a viable metabolism, and therefor the range of environments habitable by aerobic organisms, as confirmed by observations of obligately aerobic organisms in environments such as Oxygen Minimum Zones (Sperling et al. 2013b, Ulloa et al. 2013), freshwater sediment (Ettwig et al. 2010), and wastewater sludge (Yamada et al. 2006, Ward et al. 2015).

Aerobic respiration has played a critical role in the development of complex life. Aerobic respiration was a major factor in the development of eukaryotes (Katz 2012) and ultimately animals (Knoll 2011), as large, complex organisms have inherently high energetic demands that can be met by aerobic respiration but not by anaerobic metabolisms (Catling et al. 2005). While it has historically been assumed that animals required high oxygen concentrations to evolve due to inherently high O_2 demand of multicellularity (e.g. Nursall 1957), these critical oxygen concentrations have never been rigorously quantified, though they have been estimated based largely on diffusion

limitations (e.g. Runnegar 1991). Recent efforts to understand the minimum oxygen concentrations necessary to support sponges and other basal animals have demonstrated survival and growth under surprisingly low O_2 concentrations, down to about 1% PAL (Mills et al 2014, Sperling et al. 2013a). While these values are still much higher than the minimum oxygen concentrations for bacterial respiration (Stolper et al 2010), they greatly extend the spatial and temporal range of environments habitable by animals (e.g. oxygen minimum zones and oceans during the Proterozoic, e.g. Sperling et al. 2013b, Zhang et al. 2016, Fischer 2016) and are an important new constraint in thinking about oxygen as a requirement and biosignature of complex life on exoplanets

Oxygen on Mars

The history of O_2 on Mars is a curious contrast to that on Earth. While terrestrial O_2 is tied to the biological process of oxygenic photosynthesis, Mars appears to have a history of O_2 despite no apparent biological source. It is widely appreciated that the surface of Mars is, and perhaps has been since its early history, "oxidizing" (e.g., Hunten, 1979; Zent and McKay, 1994; Christensen et al., 2001; Tosca et al., 2005; McLennan et al., 2005; Goetz et al., 2005; Bibring et al., 2006; Zahnle et al., 2008; Hecht et al., 2009; Hurowitz et al. 2010; Leshin et al., 2013). Though it is the fourth most abundant species in the atmosphere, Mars currently has low absolute concentrations of O_2—a mixing ratio of 1.45×10^{-3} (Mahaffy et al. 2013). For comparison, these levels of O_2 are approximately equivalent to upper levels of O_2 allowed by geochemical proxy data on the Archean Earth (Farquhar et al. 2000; Johnson et al. 2014). However, a range of proxies suggest that oxygen concentrations may have been higher in Mars' past, perhaps when the atmosphere was thicker (e.g. Brain and

Jakosky 1998, Manga et al. 2012, Kite et al. 2014). This notion is supported by the presence of enriched manganese deposits (Lanza et al. 2014), multiple oxygen isotope ratio data collected from shallow, low-temperature carbonates preserved in the Martian meteorite ALH84001 (Farquhar et al. 1998, Halevy et al. 2011, Shaheen et al. 2015), and the presence of high concentration of oxidized salts on the surface of Mars thought to be derived from reactions involving ozone and other oxygen-derived species (e.g. Zahnle et al. 2008, Catling et al. 2010).

These results offer an interesting perspective on the habitability of Mars now and in the past. While free oxygen is essentially a trace gas in the modern Martian atmosphere, saturation of O_2 in liquid water under Martian surface conditions today would be approximately 10 nM—quite low, but within the range of oxygen concentrations shown to be utilized by *E. coli* (Stolper et al. 2010) (Figure 2). If liquid water is present on Mars (as now appears to be the case, at least transiently, e.g. Ojha et al. 2015), it could present a viable environment for aerobic microbes.

The ancient atmosphere of Mars is thought to have been much thicker than it is today, perhaps on the order of 1-2 bar (Kite et al. 2014). If the mixing ratio of O_2 were similar in this thick early atmosphere to that of modern Mars, saturation of O_2 would be on the order of ~4 μM, sufficient not only for aerobic bacteria, but in fact above the Pasteur Point and in line with the oxygen requirements of basal animals like sponges (Sperling et al. 2013a, Mills et al. 2014) (Figure 2). This is much higher than O_2 concentrations on the early Earth, perhaps reflecting relatively higher rates of hydrogen escape due to Mars' lower gravity (e.g. Lammer et al. 2013). Of course, this value is

determined utilizing an overly simple model, with a more nuanced view arising from careful analysis of paleooxygen proxies.

While many of the trace metal geochemical proxies utilized for understanding the history of oxygen on Earth either have not or cannot be applied to Mars, recent descriptions of concentrated manganese deposits in Martian rocks can help constrain ancient O_2 concentrations on Mars (Lanza et al. 2014). Manganese is the third most abundant transition metal in the crusts of Earth and Mars, where it is present only as Mn(II) and substitutes for ferrous iron in a wide variety of primary igneous minerals (Turekian and Wedepohl 1961, Taylor and McLennan, 2009). Chemical weathering of basalt, therefore, provides a substantial flux of Mn^{2+} to surface and ground waters. Mn^{2+} is soluble unless oxidized to Mn(III) or Mn(IV), which subsequently undergoes hydrolysis and forms insoluble oxide phases that rapidly sediment (Stumm and Morgan 1996). Compared to iron or sulfur, however, manganese requires uniquely high potential oxidants >500 mV (O_2 or species derived from O_2 like O_3) to undergo redox cycling and become concentrated in sedimentary rocks. In addition, the oxidation of Mn(II) is comparatively sluggish even in environments with abundant O_2, particularly in the absence of biological catalysts (Luther 2010; Morgan 2005). Due to this, Mn-rich sedimentary rocks do not appear on the Earth until after the evolution of photosynthesis and rise of oxygen (Johnson et al. 2013, Maynard 2010).

A wide range of Mn-rich materials have been discovered in sedimentary rock targets in the first 360 sols of the Curiosity rover's traverse across sedimentary rocks deposited in Gale crater (Lanza et al. 2014); these rocks contain Mn abundances that are 1–2 orders of

magnitude higher than previously observed on Mars and require redox cycling of Mn under highly oxidizing aqueous conditions. Subsequent observations made by Curiosity's instrument payload revealed that at least some of these Mn-rich phases include Mn-oxides that precipitated in an ancient groundwater aquifer (Lanza et al. 2016). These results are important because unlike the prior widespread observations of iron oxides and sulfates, Mn oxides tell us that molecular oxygen has left its mark on the early Mars geological record at a time and place in which liquid water was also present.

Oxygen on Europa

Another interesting example of worlds with a significant role played by abiotic oxygen is that of icy moons such as Europa, which may host a subsurface ocean with substantial oxygen sourced from ice radiolysis (Hand et al. 2007, Hand et al. 2009). It is generally accepted that Europa possesses an extensive subsurface ocean (Carr et al. 1998, Stevenson 2000, Hand et al. 2007), potentially containing 2-3 times the volume of water on Earth (Pappallardo et al. 1999, Stevenson 2000, Chyba and Phillips 2002). This ocean is though to be salty based on measurements of magnetism (Khurana et al., 1998; Kivelson et al., 2000), potentially dominated by magnesium sulfate (Zolotov and Shock 2001). The base of Europa's ocean is thought to be in contact with the rocky core, where water/rock interactions can provide a source of hydrogen and other reduced compounds through serpentinization-like reactions (Vance et al. 2007). These observations, as well as that of putative subduction of the icy shell (Kattenhorn and Prockter 2014), suggest a

geologically active world with sources of both reduced and oxidized compounds that could support a biosphere in the ice-covered ocean.

Observations of a tenuous oxygen atmosphere of ~10^{-11} bar were first made by Hubble Space Telescope observations (Hall et al. 1995), with the dominant source of molecular oxygen and other oxidants in the Europa system being bombardment of water and CO_2 ice on the surface by energetic electrons and ions driven by Jupiter's magnetic field (Johnson et al. 2004). As these radiolytically sourced oxidants can be trapped in ice and eventually subducted or otherwise brought into contact with Europa's ocean, this can ultimately serve as a substantial source of abiotic O_2 to any biosphere that may exist on Europa. The potential for biologically relevant oxygen accumulation in the subsurface ocean has been described previously (e.g. Chyba 2000, Hand et al. 2007). While the rate of oxygen production at the surface of Europa can be reasonably well estimated (e.g. Chyba and Phillips 2001), the rate of resurfacing and of admixture of oxygen into the subsurface ocean is not known, but will ultimately determine the flux of O_2. Estimates of resurfacing rates have been made based on crater counts and inference of subduction rates, suggesting a turnover time of <90 Ma (Zahnle et al. 2003, Kattenhorn and Prockter 2014), potentially resulting in a flux of O_2 >10^9 moles/year (Hand et al. 2007).

Depending on the rate at which Europa's ice shell is mixed with the ocean, as well as the rates at which this oxygen is consumed, oxygen may accumulate to concentrations higher than those in Earth's oceans (Hand et al. 2007). Oxygen concentrations in these environments might therefore be sufficiently high to not impose any limitations on aerobic respiration, though any net global productivity would instead likely be due to

limited fluxes of reducing power from water/rock interactions at the base of the subsurface ocean (McCollom 1999, Ward et al. 2017). Improved understanding of the oxygen requirements and metabolic efficiencies of aerobic chemolithotrophic metabolisms could allow improved predictions of biomass, which could potentially be supported on icy moons like Europa.

Role of biology in oxygen

The discovery of proxy evidence for even trace oxygen in Earth history is frequently interpreted as evidence for the presence of oxygenic Cyanobacteria (e.g. Rosing and Frei 2002, Anbar et al. 2007, Crowe et al. 2013, Planavsky et al. 2014). Meanwhile, by any number of these proxies, Mars would appear now and in the past to have been well oxygenated. However, this Martian history is certainly not interpreted as evidence for the presence of Cyanobacteria. This double standard in interpretation should serve as a reminder to be careful in assumptions about the source of oxidizing potential in the past. Oxygen need not always be biologically sourced to make an appearance in the rock record—photochemical processes can also produce oxygen and other oxidants, on the Earth as well as Mars (e.g. Haqq-Misra et al. 2011). Care must therefore be taken when interpreting evidence of oxygen, as Mars provides clear evidence that significant oxygen can be the result of purely abiotic processes. This caution must extend to the possible future interpretation of oxygen in expolanet atmospheres. Oxygen, by supporting aerobic respiration, is thought to be essential for the evolution of multicellularity and complex life, and so is an important target for detection in exoplanet atmospheres (Leger

et al. 2011, McKay 2014). While its discovery would be important for interpreting the habitability of a body, it may not necessarily reflect biological sources.

Role of oxygen in biology

While aerobic respiration is generally considered in the context of organoheterotrophy (i.e. organisms that consume fixed organic matter as both a carbon and energy source), oxygen is also utilized for biosynthesis, activation of recalcitrant metabolic substrates, as well as for aerobic respiration coupled to lithoautotrophy (i.e. organisms that oxidize inorganic compounds as an energy source to drive carbon fixation).

Oxygen is utilized in several important biosynthetic pathways, including the synthesis of sterols by eukaryotes and collagen by animals (Towe 1981). While these processes may seem minor, these and other biosynthetic pathways may account for up to 10% of O_2 consumption by mammals (Rolfe and Brown 1997). O_2 concentrations required for biosynthesis can be incredibly low, with sterol synthesis requiring O_2 at only nanomolar concentrations for the epoxidation of squalene and later oxidative demethylations of intermediates like lanosterol, but nonetheless cannot proceed under fully anoxic conditions (Waldbauer et al. 2011). The utilization of oxygen in biosynthetic processes appears to be a relatively late innovation, after the radiation of the three domains of life (Raymond and Segre 2006). This suggests that while molecular oxygen may not be necessary for the origin of life, it does provide an opportunity for significant diversification of biochemical processes and may play a role in the development of

complex life beyond just serving as a highly energetic electron acceptor (Raymond and Segre 2006, Falkowski 2006).

O_2 is furthermore essential for the activation of otherwise inaccessible electron donors. Compounds such as methane and ammonia must first be activated utilizing O_2 via a monooxygenase enzyme in a non-energy conserving step before they can subsequently be further oxidized to conserve energy (Hanson and Hanson 1996). While these processes have an absolute requirement for oxygen, they can make use of exceedingly low oxygen concentrations. Half saturation constants for ammonia oxidation are on the order of ~300 nM O_2 (Bristow et al. 2016), and both methane and ammonia oxidation has been observed in environments with much less than 1 μM O_2 (Hatzenpichler 2012, Kits et al. 2015, Bristow et al. 2016).

Finally, oxygen is utilized as an electron acceptor in many lithotrophic metabolisms. Oxygen reduction can be bioenergetically coupled to the oxidation of reduced sulfur compounds (such as sulfide and elemental sulfur), ferrous iron, molecular hydrogen, and reduced nitrogen compounds (ammonium and nitrite). These processes are carried out by diverse bacteria and archaea, and are essential for much of the biogeochemical nutrient cycles on Earth. The reduction of oxygen in these processes is through a combination of respiratory electron transfer through the HCOs and *bd* oxidase in addition to utilization of molecular oxygen to activate recalcitrant substrates (such as the activation of ammonium to hydroxylamine by ammonia monooxygenase). These various enzymes may have different oxygen requirements; further characterization of the oxygen requirements of the oxygenase enzymes in the context of lithoautotrophs will

help to constrain what if any of these metabolic couples may be viable in modern or ancient Martian environments.

Because the oxidation of reduced compounds by O_2 is often a competition between the spontaneous abiotic reaction and microbes performing this process for energy conservation, microbes have adapted to microaerobic conditions where the abiotic rate is sufficiently sluggish. Iron oxidizing bacteria, for instance, typically grow under microaerobic conditions below 50 μM O_2, and have been demonstrated to grow at O_2 concentrations as low as 3 μM (Chan et al. 2016). Whether this is an ecological lower limit or a biochemical one is unclear, as the thermodynamics of iron oxidation remain favorable even under much lower O_2 concentrations (Emerson et al. 2010). Furthermore, the yield of carbon fixed by these metabolisms is poorly constrained, but appears to be significantly less efficient than ideal stoichiometries (e.g. Neubauer et al. 2002, Ward et al. 2017), adding additional uncertainty to the potential extent of aerobic lithotrophic metabolisms on other worlds.

Potentially viable metabolisms on Mars and other worlds

The mounting evidence for liquid water throughout Mars' past as well as potentially on the modern surface (Martin-Torres et al. 2015, Grotzinger et al. 2015, Ojha et al. 2015), combined with the apparent availability of free oxygen, suggests a history of Martian environments that are not only habitable but also present a variety of thermodynamically favorable metabolic couples. For instance, the putative discovery of

methane on Mars (Formisano et al 2004, Mumma et al. 2009, Webster et al. 2015) along with metabolically useful oxygen presents the opportunity for aerobic methanotrophy, the metabolic oxidation of methane to CO_2 using O_2 while conserving energy, a process driven by diverse bacteria on Earth (e.g. Hanson and Hanson 1996). This metabolism requires molecular oxygen for the activation of methane, and typically couples this process to aerobic respiration (Hanson and Hanson 1996). Recently, however, it has been discovered that the terminal electron accepting process for this metabolism is more versatile than previously realized, with the potential in some bacteria to couple aerobic methanotrophy to nitrate reduction when oxygen concentrations are low (e.g. Kits et al. 2015, Skennerton et al. 2015). The recent discovery of nitrate in Martian sediment (Stern et al. 2015) therefore supports the potential viability of this metabolism in some Martian environments as well.

Other potentially viable metabolisms on the Martian surface include the aerobic oxidation of ferrous iron. Iron oxidizing bacteria are typically microaerobic, with preferred oxygen concentrations on the order of ~1-100 μM O_2, but likely functioning down to much lower concentrations (Chan et al. 2016). As the metabolic byproduct of iron oxidation is solid ferric iron oxides, iron oxidizing bacteria must shed ferric minerals, typically in diagnostic, preservable morphologies such as sheathes and stalks (e.g. Chan et al. 2011). If similar microbes exist on Mars, iron biominerals could be observable biosignatures. The yield of carbon fixed per iron oxidized by iron oxidizing bacteria is poorly constrained: the ideal electron-balanced stoichiometry of this metabolism is 4 Fe: 1C, yet experimental yields are much poorer, on the order of 40 Fe:

1C (Neubauer et al. 2002), with similarly poor yields in natural systems (e.g. Ward et al. 2017). As a result, the size of biosphere and amount of productivity this metabolism could support is poorly constrained.

Meanwhile, Europa is predicted to have extensive water-rock interactions at the base of the ocean, potentially supporting ecosystems analogous to hydrothermal vents on Earth. The geochemical makeup of these systems on Europa is not well constrained, but likely supply a flux of reduced compounds such as molecular hydrogen, ferrous iron, and/or sulfide. These compounds can all be aerobically oxidized by lithoautotrophs, potentially supporting primary productivity entirely detached from photosynthesis.

The interaction of reduced compounds from water/rock interactions at Europa's core and oxygen sourced from the icy shell could support aerobic chemoautotrophic metabolisms, such as hydrogen, sulfur, or ferrous iron oxidation. Biomass estimates have been derived for Europa (e.g. Chyba and Phillips 2001, Zolotov and Shock 2003, McCollom 1999, Hand et al. 2009); these estimates depend in part on the limits of aerobic respiration, and the yield of lithotrophic metabolisms. Both of these factors are only poorly constrained (e.g. Chan et al. 2016, Ward et al. 2017, Neubauer et al. 2002). As a result, the potential for the origin and survival of a substantial biosphere is contested (e.g. Gaidos et al. 1999). However, the depth of rock fracturing and therefore water-rock interactions on Europa may be much deeper than on Earth (Vance et al. 2007), potentially increasing the flux of reduced compounds available for biological productivity.

Oxygen fluxes into Europa's ocean may be quite high, particularly given recent evidence for subduction indicating resurfacing at <90 Ma timescales (Kattenhorn and

Prockter 2014). The resulting rates of delivery of oxygen from radiolysis on the surface into the ocean may even result in dissolved O_2 in Europa's ocean on the order of 10^{-3} moles/liter, nearly an order of magnitude higher than saturation at the surface of Earth's ocean today (Hand et al. 2007) (Figure 2). Recent analysis of the redox balance of Europa suggests that hydrogen fluxes from serpentinization-like reactions are comparable to oxygen delivery from the surface (Vance et al. 2016). Europa could therefore support a biosphere at a stable redoxicline at some region between the base and top of the ocean—the exact location and extent of this productive redoxicline could have significant bearing on life detection on Europa, and whether this could be viable from a plume-sampling flyby or if direct sampling of the Europan ocean would be necessary.

Evolutionary history of respiration

One caveat to the viability of aerobic respiration on Mars today is that while aerobic respiration continues down to vanishingly low oxygen concentrations, it may not be sufficient for supporting growth. Aerobic respiration is accomplished using a complex electron transport chain, which culminates in the enzymatic reduction of O_2 to H_2O, a step that can be accomplished using any one of several proteins with disparate evolutionary histories. These include the heme copper O_2 reductase superfamily (A-, B-, and C-families, which are all related), and the *bd* oxidases, which evolved independently (Borisov et al 2011). The A-family heme copper O_2 reductases (HCO) are the most widespread in nature (including in eukaryotic mitochondria), and require the highest O_2 levels to be effective, but also conserve the most energy (Han et al 2011). The other families (including the *bd* oxidases) conserve less energy but are effective at lower O_2

concentrations (Morris and Schmidt 2013). While it is unclear whether this inverse correlation between energy conservation and O_2 affinity is the result of biochemical optimization or evolutionary contingency, it must be considered that respiration under relatively low oxygen concentrations, such as those that may have existed on early Mars, may not be capable of supporting the same amount of biomass as respiration at higher oxygen concentrations. Further investigation into the relationship of O_2 concentrations, energy conservation, and the evolutionary history of aerobic respiration can help to resolve this uncertainty.

Oxygen and the evolution of complex life

The possibility of oxygen accumulating to substantial concentrations through purely abiotic processes has exciting implications for the evolution of life on other planets. Oxygenic photosynthesis is an evolutionary singularity, and has evolved only once in Earth history through the complicated coupling of two independent photosystems (Falkowski 2011). This innovation was essential for the eventual evolution of complex life on Earth, but is the unlikely result of evolutionary contingency. The evolution of complex multicellularity, on the other hand, has occurred independently several times on Earth (e.g. in the animals, fungi, plants, and red algae)(Knoll et al. 2011), and simple forms of multicellularity have even been artificially evolved over short timescales in the laboratory (Ratcliff et al. 2012). The evolution of multicellularity may therefore be "easier", though it has only occurred on Earth after the rise of oxygen. It is therefore possible that on other worlds where oxygen accumulates through abiotic processes,

multicellularity could evolve more quickly and easily, as it does not require waiting for the "hard step" of oxygenic photosynthesis (Figure 3).

A potential caveat to abiotic oxygen improving the odds of evolution of complex life is the impact of oxidizing conditions on the origin of life. While conditions necessary for the origin of life are not known, most proposed environments for the origin of life, such as hydrothermal vents, are characterized by reducing conditions (e.g. Baross and Hoffman 1985, Martin et al. 2008). It is therefore possible that oxidizing conditions as a result of abiotic oxygen as proposed here could inhibit the origin of life, or at least restrict the environments in which it could occur. Abiotic oxygen could therefore be an evolutionary Catch 22: it can jumpstart the evolution of complex life, but prevents life from originating to begin with. This impasse could be solved by panspermia, in which life may originate on a world more favorable for its origin and then be transferred to one more conducive to its growth and evolution. While panspermia is usually discussed in the context of life originating elsewhere and subsequently seeding Earth (e.g. Kirschvink and Weiss 2002), this process could occur in either direction (e.g. Melosh 1988). As a result, life may have the opportunity to originate in one environment and subsequently thrive in another, at least in solar systems such as our own with a range of planetary environments.

Conclusions

The history of oxygen on Earth is, fundamentally, a biological one. O_2 is produced in staggering quantities as a result of oxygenic photosynthesis by Cyanobacteria, and is rapidly recycled as a result of aerobic respiration. This state of affairs has continued for the past 2.3 Ga, and has supported the radiation of many

microbial groups as well as the origin and diversification of eukaryotes, and, eventually, complex multicellular organisms like plants, animals, and fungi. Before the rise of oxygenic photosynthesis, the Earth was essentially anoxic and supported only very limited productivity, and all life was microbial. The development of complex life therefore required multiple preconditions, some of which appear to have been very challenging for biology to evolve. Among these, oxygenic photosynthesis may be one of the hardest steps. This metabolism is biochemically complicated, required evolutionary contingencies to diverge and recombine two photosystems, and has evolved only a single time in Earth history. In comparison, aerobic respiration has multiple independent origins (heme copper oxidases, alternative oxidases, and *bd* oxidases), and multicellularity has even more (e.g. animals, plants, fungi, brown algae and red algae). The relatively late evolution of oxygenic photosynthesis may therefore have delayed the origin of complex life on Earth for a billion years or more. If oxygen can be produced abiotically in large quantities on other planets, then aerobic respiration –and the energetic, complex organisms that depend on it—may be capable of evolving more quickly. The presence of oxygen in a planet's atmosphere or oceans may therefore be an important indicator of habitability, as even abiotic oxygen may provide a path to jumpstarting the emergence of complex life. We therefore propose a strategy of "Follow the Oxygen", in which environments with opportunities for energetic metabolisms such as aerobic respiration are targeted for investigation and life detection. If life has evolved in these environments, it may be more likely to have developed into more complex forms than on Earth-like

worlds where oxygenic photosynthesis was a necessary precondition for highly energetic organisms.

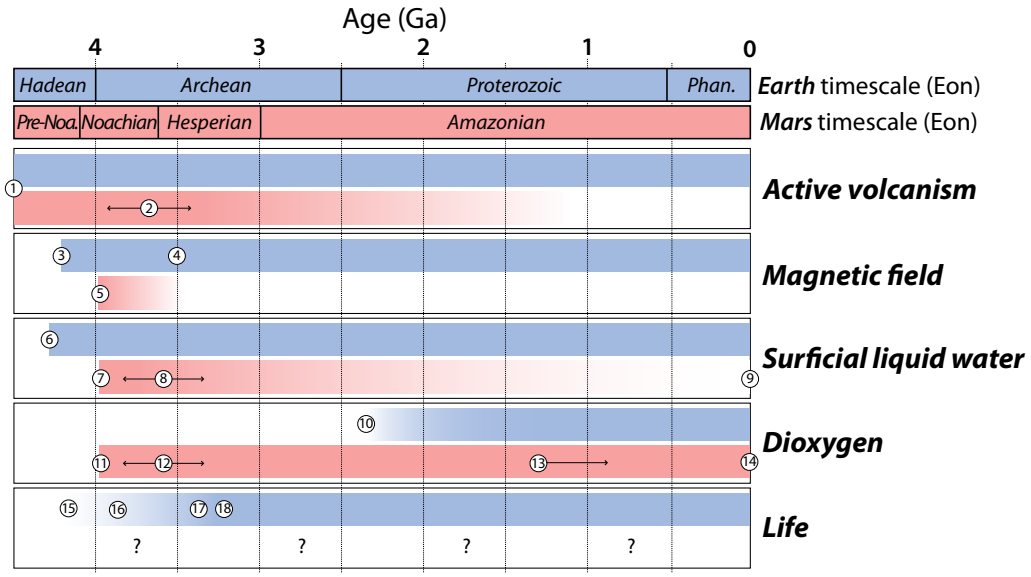


Figure 1: Comparative timelines of Earth and Mars. While the surfaces of Earth and Mars are now quite different, their early histories were more similar, with early Mars having once had a magnetic field, surficial water, active volcanism, and a dense atmosphere similar to Earth. Mars, however, appears to have had significant O_2 and ozone early in its history, in contrast to Earth.

1 - Active volcanism assumed due to the thermal status of both planets early in their history.

2 - Tharsis volcanic province (Bouley et al. 2016).

3 - Paleomagnetic data from Jack Hills zircons (Tarduno et al. 2015).

4 - Paleomagnetic data from early Archean greenstone belts (Tarduno et al. 2010, Biggin et al. 2011)

5 - Magnetic studies of ALH84001 (Weiss et al. 2002).

6 - Elevated O isotope ratios in Jack Hills zircons (Valley et al. 2002).

7 - Clumped isotope ratios of carbonates in ALH84001 (Halevy et al. 2011).

8 - Fluvio-deltaic strata in Gale Crater (Grotzinger et al. 2015) and Aeolis Dorsa (DiBiase et al. 2013).

9 - Hydrated salts associated with recurring slope lineae (Ojha et al. 2015).

10 - Great oxygenation event (Fischer et al. 2016).

11 - 17O mass anomalies in low-temperature carbonate salts in ALH84001 (Farquhar et al. 1998).

12 - Mn enrichments (Lanza et al. 2014) and high valent Cl-bearing salts (Leshin et al. 2013, Farley et al. 2016) in fluvio-deltaic strata deposited in Gale Crater.

13 - 17O mass anomalies in carbonate and sulfate salts of Nakhla and Lafayette (Farquhar et al. 2000).

14 - O2 mixing ratio of 1.45(+/-0.09)x10^-3 (by volume) in modern Mars atmosphere in Gale crater from Curiosity's quadrupole mass spectrometer (Mahaffy et al. 2013).

15 - Earliest speculative evidence for life from graphitic carbon inclusion in Jack Hills zircons (Bell et al. 2015).

16 - Promising but non-unique evidence for life graphite-bearing turbidite metasedimentary rocks from the Isua Supracrustal Belt (Rosing 1997).

17 - Earliest likely microfossils in sedimentary chert of the Strelley Pool Formation (Sugitani et al. 2010).

18 - Oldest certain microfossils in fine-grained clastics of the Moodies Formation (Javaux et al. 2010).

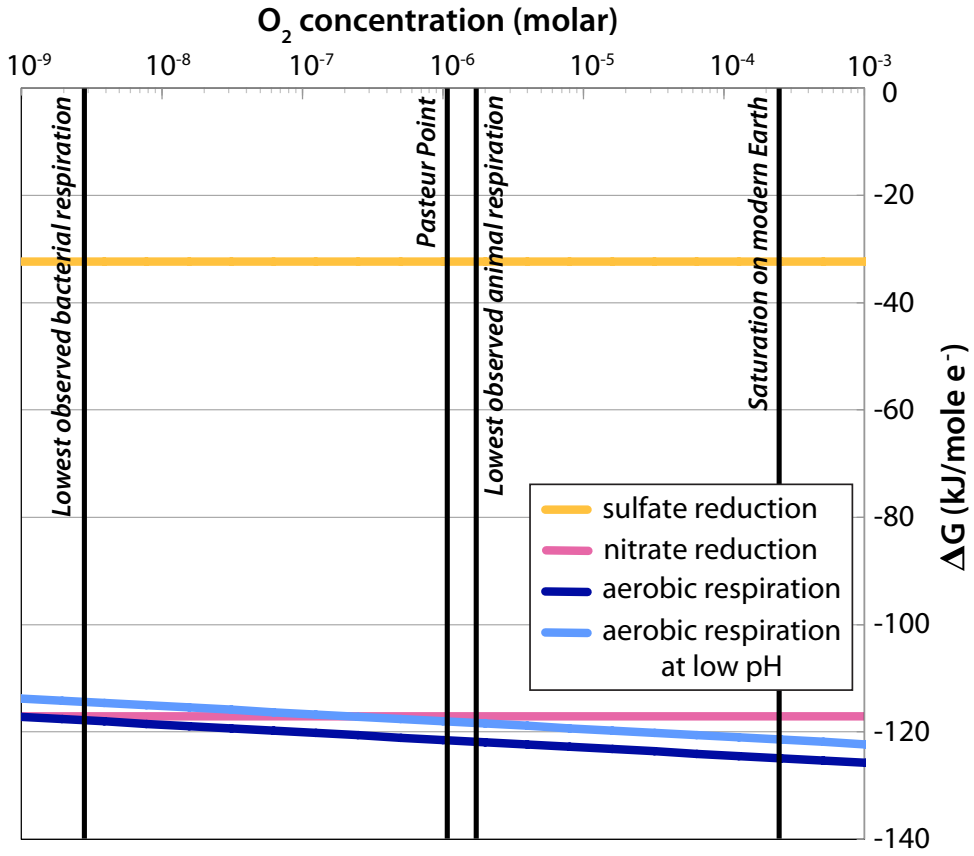


Figure 2: Thermodynamic favorability of microbial metabolisms as Gibbs free energy of reaction across a range of oxygen concentrations. Even at very low O_2 concentrations, aerobic respiration is still highly favorable ($\Delta G \ll 0$). Also plotted are lines denoting relevant oxygen concentrations, including known limits of respiration for bacteria (Stolper et al. 2010) and animals (Mills et al. 2014), as well as saturation under surface atmospheric conditions on modern Earth and Mars. Even the thin modern Martian atmosphere contains sufficient oxygen to support respiration by bacteria such as *E. coli*.

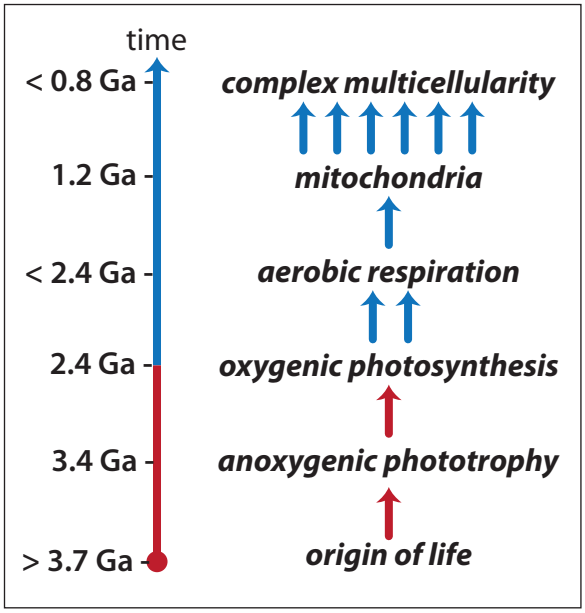


Figure 3: Evolutionary contingency and the origin of complex life. Several steps had to occur for complex life to evolve on Earth. Beginning with the origin of life, evolution of anoxygenic phototrophy, coupling in series of two reaction to perform oxygenic photosynthesis, evolution of aerobic respiration, evolution of eukaryotes and endosymbiosis of the mitochondrion, and finally the evolution of complex multicellularity. Of these, some steps have occurred independently multiple times, while others are evolutionary singularities. On a world with a significant abiotic O_2 flux, life may be able to skip the hard step of evolving oxygenic photosynthesis and jumpstart the evolution of complexity. Arvidson, RE et al. High concentrations of manganese and sulfur in deposits on Murrary Ridge, Endeavour Crater, Mars. American Mineralogist, in press.

References:

1. Bargar JR, Tebo BM, Bergmann U, Webb SM, Glatzel P, Chiu VQ, Villalobos M (2005) Biotic and abiotic products of Mn (II) oxidation by spores of the marine Bacillus sp. strain SG-1. American Mineralogist, 90, 143-154.
2. Baross, J.A. & Hoffman, S.E. Origins Life Evol Biosphere (1985) 15: 327.
3. Bekker A, Holland HD, Wang PL, Rumble D, Stein HJ, Hannah JL, Coetzee LL, Beukes NJ (2004) Dating the rise of atmospheric oxygen. Nature, 427:117-120.
4. Berkner LV, Marshall LC (1965) On the origin and rise of oxygen concentration in the Earth's atmosphere. J. Atmos. Sci., 22, 225-261.
5. Bibring JP et al. (2006) Global mineralogical and aqueous Mars history derived from OMEGA/Mars Express data. Science, 312, 400–404.
6. Borda, M. J., Elsetinow, A. R., Schoonen, M. A. & Strongin, D. R. Pyrite-Induced Hydrogen Peroxide Formation as a Driving Force in the Evolution of Photosynthetic Organisms on an Early Earth. *Astrobiology* **1,** 283–288 (2001).
7. Borisov VB, Gennis RB, Hemp J, Verkhovsky MI (2011) The cytochrome bd respiratory oxygen reductases. Biochimica et Biophysica Acta-Bioenergetics, 1807, 1398-1413.
8. Bouley, S. *et al*. Late Tharsis formation and implications for early Mars. *Nature* **531,** 344–347 (2016).

9. Brain, D. a. & Jakosky, B. M. Atmospheric loss since the onset of the Martian geologic record: Combined role of impact erosion and sputtering. *J. Geophys. Res.* **103,** 22689 (1998).
10. Bristow, L. A. *et al.* Ammonium and nitrite oxidation at nanomolar oxygen concentrations in oxygen minimum zone waters. *Proc. Natl. Acad. Sci. U. S. A.* **113,** 10601–6 (2016).
11. Carr, Michael H., et al. "Evidence for a subsurface ocean on Europa." *Nature*391.6665 (1998): 363-365.
12. Catling, D. C. & Claire, M. W. How Earth's atmosphere evolved to an oxic state: A status report. *Earth Planet. Sci. Lett.* **237,** 1–20 (2005).
13. Catling, D. C., Glein, C. R., Zahnle, K. J. & McKay, C. P. Why O2 is required by complex life on habitable planets and the concept of planetary 'oxygenation time'. *Astrobiology* **8,** 377–395 (2006).
14. Catling, D. C. *et al.* Atmospheric origins of perchlorate on mars and in the atacama. *J. Geophys. Res. E Planets* **115,** 1–15 (2010).
15. Chan CS, Fakra SC, Emerson D, Fleming EJ, Edwards KJ (2011) Lithotrophic iron-oxidizing bacteria produce organic stalks to control mineral growth: implications for biosignature formation. ISME Journal 5, 717–727.
16. Chan, C., Emerson, D. & Luther, G. W. The role of microaerophilic Fe-oxidizing microorganisms in producing banded iron formations. *Geobiology* (2016).
17. Chyba, C. F. Energy for microbial life on Europa. *Nature* **403,** 381–382 (2000).

18. Chyba, C. F. & Phillips, C. B. Possible ecosystems and the search for life on Europa. *Proc. Natl. Acad. Sci.* **98,** 801–804 (2001).

19. Christensen PR, Morris RV, Lane MD, Bandfield JL, Malin MC (2001) Global mapping of Martian hematite mineral deposits: Remnants of water-driven processes on early Mars. J. Geophys. Res., 106(E10), 23,873–23,885.

20. Cockell, C. S., et al. "Habitability: A Review." *Astrobiology* 16.1 (2016): 89-117

21. Crowe, S. A. *et al*. Atmospheric oxygenation three billion years ago (SI). *Nature* **501,** 535–538 (2013).

22. Ehlmann, B. L. *et al*. Subsurface water and clay mineral formation during the early history of Mars. *Nature* **479,** 53–60 (2011).

23. Emerson, D., Fleming, E. J. & McBeth, J. M. Iron-oxidizing bacteria: an environmental and genomic perspective. *Annu. Rev. Microbiol.* **64,** 561–583 (2010).

24. Ettwig, K. F. *et al*. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* **464,** 543–548 (2010).

25. Falkowski, Paul G. "Tracing oxygen's imprint on Earth's metabolic evolution."*Science* 311.5768 (2006): 1724-1725.

26. Falkowski, PG. The biological and geological contingencies for the rise of oxygen on earth. *Photosynth. Res.* **107,** 7–10 (2011).

27. Farley, K. A., et al. "Light and variable 37 Cl/35 Cl ratios in rocks from Gale Crater, Mars: Possible signature of perchlorate." *Earth and Planetary Science Letters* 438 (2016): 14-24.

28. Farquhar J et al. (1998) Atmosphere-surface interactions on Mars: Δ17O measurements of carbonate from ALH 84001. Science 280:1580–1582.

29. Farquhar J, Bao H, Thiemens M. (2000) Atmospheric influence of Earth's earliest sulfur cycle. Science, 289:756-758.

30. Fischer, W. W. Breathing room for early animals. *PNAS* (2016). doi:10.1038/382111a0

31. Fischer, W., Hemp, J. & Johnson, J. E. Evolution of Oxygenic Photosynthesis. *Annu. Rev. Earth Planet. Sci.* **44,** (2016).

32. Formisano, V., Atreya, S., Encrenaz, T., Ignatiev, N. & Giuranna, M. Detection of methane in the atmosphere of Mars. *Science (80-.).* **306,** 1758–1761 (2004).

33. Fassett, C. I. & Head, J. W. Sequence and timing of conditions on early Mars. *Icarus* **211,** 1204–1214 (2011).

34. Gaidos, E. J., Nealson, K. H. & Kirschvink, J. L. Life in Ice-Covered Oceans. *Science (80-.).* **284,** 1631–1633 (1999).

35. Goetz W et al. (2005) Indication of drier periods on Mars from the chemistry and mineralogy of atmospheric dust. Nature, 436, 62–65.

36. Grotzinger, J. P. *et al*. Deposition, exhumation, and paleoclimate of an ancient lake deposit, Gale crater, Mars. *Science (80-.).* **350,** aac7575–aac7575 (2015).

37. Halevy I et al. (2011) Carbonates in the Martian meteorite Allan Hills 84001 formed at 18 ± 4 degrees C in a near-surface aqueous environment, Proceedings of the National Academy of Sciences, 108, 16895-16899

38. Hall, D. T., Strobel, D. F., Feldman, P. D., McGrath, M. A. & Weaver, H. A. Detection of an oxygen atmosphere on Jupiter's moon Europa. *Nature* **373,** 677–681 (1995).

39. Han H, Hemp J, Pace LA, Ouyang H, Ganesan K, Roh JH, Daldal F, Blanke SR, Gennis RB (2011) Adaptation of aerobic respiration to low O2 environments. PNAS, 108,14109-14114.

40. Hand KP, Brown ME (2013) Keck II Observations of Hemispherical Differences in H2O2 on Europa. The Astrophysical Journal Letters, 766:2, L21, doi: 10.1088/2041-8205/766/2/L21.

41. Hand KP et al. (2007) Energy, chemical disequilibrium and geological constraints on Europa. Astrobiology, 7, 1006-1022.

42. Hand KP, Chyba CF, Priscu JC, Carlson RW, Nealson KH (2009) Astrobiology and the Potential for Life on Europa. In: Europa, edited by RT Pappalardo, WB McKinnon, and K Khurana, University of Arizona Press, Tucson, AZ.

43. Hanson, R. S., and Hanson, T. E. (1996). Methanotrophic bacteria. *Microbiol. Rev*. 60, 439–471.

44. Haqq-Misra, J., Kasting, J. F. & Lee, S. Availability of O(2) and H(2)O(2) on pre-photosynthetic Earth. *Astrobiology* **11,** 293–302 (2011).

45. Hatzenpichler, R. Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol*. **78,** 7501–7510 (2012).

46. Hecht MH et al. (2009) Detection of perchlorate and the soluble chemistry of Martian soil at the Phoenix Lander site. Science, 325, 64–67.

47. Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press, Princeton, NJ.

48. Hunten DM (1979) Possible oxidant sources in the atmosphere and surface of Mars. J. Mol. Evol., 14, 71–78.

49. Hurowitz JA et al. (2010) Origin of acidic surface waters and the evolution of atmospheric chemistry on early Mars. Nature Geoscience, 3, 323-326.

50. Jakosky, B. M. & Phillips, R. J. Mars' volatile and climate history. *Nature* **412,** 237–244 (2001).

51. Johnson, R.E., Carlson, R.W., Cooper, J.F., Paranicas, C., Moore, M.H., and Wong, M.C. (2004) Radiation effects on the surface of the Galilean satellites. In Jupiter, edited by F. Bagenal, T.E. Dowling, and W.B. McKinnon, Cam- bridge University Press, Cambridge, pp. 485–512.

52. Johnson JE et al. (2013) Manganese-oxidizing photosynthesis before the rise of cyanobacteria. PNAS, 110, 11238-11243.

53. Johnson JE et al. (2014) O2 constraints from Paleoproterozoic detrital pyrite and uraninite. Geological Society of America Bulletin, Published online ahead of print on 27 Feb. 2014, doi: 10.1130/B30949.1.

54. Kattenhorn, S. A. & Prockter, L. M. Evidence for subduction in the ice shell of Europa. *Nat. Geosci.* **7,** 762–767 (2014).

55. Katz LA (2012) Origin and diversification of eukaryotes. Annu. Rev. Micriobiol., 66, 411-427.

56. Kirschvink, Joseph L., and Benjamin P. Weiss. "Mars, panspermia, and the origin of life: where did it all begin." *Palaeontologia electronica* 4.2 (2002): 8-15.

57. Kite ES et al. (2014) Low palaeopressure oft he Martian atmosphere estimated from the size distribution of ancient craters. Nature Geoscience, 7, 335-339.

58. Kits, K. D., Klotz, M. G., and Stein, L. Y. (2015). Methane oxidation coupled to nitrate reduction under hypoxia by the Gammaproteobacterium Methylomonas denitrificans, sp. nov. type strain FJG1. *Environ. Microbiol.*

59. Knoll AH (2011) The multiple origins of complex multicellularity. Annual Reviews of Earth and Planetary Sciences, 39, 217-239.

60. Lammer, H. *et al*. Outgassing history and escape of the Martian atmosphere and water inventory. *Space Sci. Rev*. **174,** 113–154 (2013).

61. Lanza NL et al. (2014) High manganese concentrations in rocks at Gale crater, Mars. Geophysical Research Letters, 41.

62. Lanza, Nina L., et al. "Oxidation of manganese in an ancient aquifer, Kimberley formation, Gale crater, Mars." *Geophysical Research Letters* 43.14 (2016): 7398-7407.

63. Léger, a *et al*. Is the presence of oxygen on an exoplanet a reliable biosignature? *Astrobiology* **11,** 335–341 (2011).

64. Leshin LA et al. (2013) Volatile, isotope, and organic analysis of Martian fines with the Mars Curiosity rover. Science, 341, 1238937-1-9.

65. Liang, M.-C., Hartman, H., Kopp, R. E., Kirschvink, J. L. & Yung, Y. L. Production of hydrogen peroxide in the atmosphere of a Snowball Earth and the

origin of oxygenic photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 18896–18899 (2006).

66. Lu Z, Chung Chang Y, Yin QZ, Ng CY, Jackson WM (2014) Evidence for direct molecular oxygen production in CO2 photodissociation. Science, 346, 61.
67. Luther GW (2010) The role of one- and two-electron transfer reactions in forming thermodynamically unstable intermediates as barriers in multi-electron redox reactions. Aquat. Geochem., 16, 395–420.
68. Lyons TW, Reinhard CT, Planavsky NJ (2014) The rise of oxygen in Earth's early ocean and atmosphere. Nature, 506, 307-315.
69. Mahaffy PR et al. (2013) Abundance and isotopic composition of gases in the Martian atmosphere from the Curiosity rover. Science, 341, 263-266.
70. Manga, M., Patel, A., Dufek, J. & Kite, E. S. Wet surface and dense atmosphere on early Mars suggested by the bomb sag at Home Plate, Mars. *Geophys. Res. Lett.* **39,** 2–6 (2012).
71. Martin, William, et al. "Hydrothermal vents and the origin of life." *Nature Reviews Microbiology* 6.11 (2008): 805-814.
72. Martín-Torres, F. J. *et al.* Transient liquid water and water activity at Gale crater on Mars. *Nat. Geosci.* **8,** 1–5 (2015).
73. Maynard JB (2010) The chemistry of manganese ores through time: a signal of increasing diversity of Earth-surface environments. Economic Geology, 105:535-552.
74. McCollom TM (1999) Methanogenesis as a potential source of chemical energy

for primary biomass production by autotrophic organisms in hydrothermal systems on Europa. Journal of Geophysical Research, 104, 30729-30742.

75. McKay, C. P. Requirements and limits for life in the context of exoplanets. *Proc. Natl. Acad. Sci.* **111,** 12628–12633 (2014).
76. McLennan SM et al. (2005) Provenance and diagenesis of the evaporate-bearing Burns formation, Meridiani Planum, Mars. Earth Planet. Sci. Lett., 240, 95–121.
77. Melosh, H.J. (1988) The rocky road to Panspermia. Nature 332, 687–688.
78. Mills DB, Ward LM, Jones CA, Sweeten B, Forth M, Treutsch AH, Canfield DE (2014) Oxygen requirements of the earliest animals. PNAS, 111, 4168-4172.
79. Morgan JJ (2005) Kinetics of reaction between O2 and Mn(II) species in aqueous solutions. Geochimica et Cosmochimica Acta, 69, 35-48.
80. Morris RL, Schmidt TM (2013) Shallow breathing: bacterial life at low O2. Nature Reviews Microbiology, 11, 205-212.
81. Mumma, Michael J, Villanueva, G.L., Novak, R.E., Hewagama, T., Bonev, B.P., DiSanti, M.A., Mandell, A.M., Smith, M. D. Strong Release of Methane on Mars in Northern Summer 2003. *Science (80-.).* **323,** 1041–1045 (2009).
82. Neubauer, S. C. *et al.* Life at the energetic edge: kinetics of circumneutral Fe oxidation by lithotrophic iron oxidizing bacteria isolated from the wetland plant rhizosphere. *Appl. Environ. Microbiol.* **68,** 3988–3995 (2002).
83. Ojha, L. *et al.* Spectral evidence for hydrated salts in recurring slope lineae on Mars. *Nat. Geosci.* (2015). doi:10.1038/ngeo2546

84. Pappalardo, R. T. et al.: 1999, 'Does Europa have a subsurface ocean? Evaluation of the geological evidence', J. Geophys. Res. 104, 24,015–24,055.

85. Pavlov, Alexander A., Lisa L. Brown, and James F. Kasting. "UV shielding of NH3 and O2 by organic hazes in the Archean atmosphere." *Journal of Geophysical Research: Planets* 106.E10 (2001): 23267-23287.

86. Ratcliff, W. C., Denison, R. F., Borrello, M. & Travisano, M. Experimental evolution of multicellularity. *Proc. Natl. Acad. Sci.* **109,** 1595–600 (2012).

87. Raymond, J. & Segre, D. The effect of oxygen on biochemical networks and the evolution of complex life. *J. Chem. Inf. Model.* **311,** 1764–1767 (2006).

88. Rolfe, D. F. & Brown, G. C. Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiol. Rev.* **77,** 731–758 (1997).

89. Runnegar, B. Precambrian oxygen levels estimated from the biochemistry and physiology of early eukaryotes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **97,** 97–111 (1991).

90. Rye, R. & Holland, H. D. Paleosols and the evolution of atmospheric oxygen: A critical review. *Am. J. Sci.* **298,** 621–672 (1998).

91. Shaheen R et al. (2015) Carbonate formation events in ALH 84001 trace the evolution of the Martian atmosphere. PNAS, 112, 336-341.

92. Shih, PM. Cyanobacterial Evolution: Fresh Insight into Ancient Questions. *Curr. Biol.* **25,** R192–R193 (2015).

93. Shih, P., Hemp, J., Ward, L., Matzke, N. & Fischer, W. Crown group oxyphotobacteria postdate the rise of oxygen. *Geobiology* (2016).

94. Skennerton, CT, LM Ward, A Michel, et al. 2015. Genomic reconstruction of an uncultured hydrothermal vent gammaproteobacterial methanotroph (family Methylothermaceae) indicates multiple adaptations to oxygen limitation. Frontiers in Microbiology 6 1425.

95. Sperling, E. A., Halverson, G. P., Knoll, A. H., Macdonald, F. A. & Johnston, D. T. A basin redox transect at the dawn of animal life. *Earth Planet. Sci. Lett.* **371-372,** 143–155 (2013a).

96. Sperling, E. A. *et al*. Oxygen, ecology, and the Cambrian radiation of animals. *Proc. Natl. Acad. Sci.* **110,** 13451–13466 (2013b).

97. Stern, J. C. *et al*. Evidence for indigenous nitrogen in sedimentary and aeolian deposits from the Curiosity rover investigations at Gale crater , Mars. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 4245–4250 (2015).

98. Stevenson, David. "Europa's ocean--the case strengthens." *Science*289.5483 (2000): 1305-1307.

99. Stolper DA, Revsbech NP, Canfield DE (2010) Aerobic growth at nanomolar oxygen concentrations. PNAS, 107, 18755-18760.

100. Stumm W, Morgan JJ (1996) Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters, 3rd ed., 1042 pp., John Wiley, New York.

101. Sugitani, K. *et al*. Biogenicity of morphologically diverse carbonaceous microstructures from the ca. 3400 Ma Strelley pool formation, in the Pilbara Craton, Western Australia. *Astrobiology* **10,** 899–920 (2010).

102. Tebo BM, Bargar JR, Clement BG, Dick GJ, Murray KJ, Parker D, Verity R,

Webb SM (2004) Biogenic manganese oxides: properties and mechanisms of formation. Ann. Rev. Earth Planet. Sci., 32, 287–328.

103. Tosca NJ, McLennan SM, Clark BC, Grotzinger JP, Hurowitz JA, Knoll AH, Schröder C, Squyres SW (2005) Geochemical modeling of evaporation processes on Mars: Insight from the sedimentary record at Meridiani Planum. Earth Planet. Sci. Lett., 240, 122–148.

104. Towe, K.M. (1981) Biochemical keys to the emergence of complex life. In Life in the Universe, edited by J. Billingham, MIT Press, Cambridge, MA, pp. 297–306

105. Turekian KK, Wedepohl KH (1961) Distribution of the elements in some major units of the Earth's crust. Geol. Soc. Am. Bull., 72(2), 175–192.

106. Ulloa, Osvaldo, et al. "Microbial oceanography of anoxic oxygen minimum zones." *Proceedings of the National Academy of Sciences* 109.40 (2012): 15996-16003.

107. Vance, S. *et al*. Hydrothermal systems in small ocean planets. *Astrobiology* **7,** 987–1005 (2007).

108. Vance, S. D., K. P. Hand, and R. T. Pappalardo. "Geophysical controls of chemical disequilibria in Europa." *Geophysical Research Letters* (2016).

109. Waldbauer, J. R., NEWMAN, D. K. & Summons, R. E. Microaerobic steroid biosynthesis and the molecular fossil record of Archean life. *Proc. Natl. Acad. Sci*. **108,** 13409–13414 (2011).

110. Ward, Lewis M., et al. "Draft genome sequence of Leptolinea tardivitalis YMTK-2, a mesophilic anaerobe from the Chloroflexi class Anaerolineae."*Genome announcements* 3.6 (2015): e01356-15.

111. Ward, LM, Kirschvink, JL, Fischer, WW (2016) Timescales of Oxygenation Following the Evolution of Oxygenic Photosynthesis. *Orig. Life Evol. Biosph.* 46(1) pp51-65.

112. Ward, LM et al. Microbial Diversity of Okuoku-hachikurou Onsen, a Japanese hot spring analog of Banded Iron Formations. Geobiology, submitted.

113. Webster, C. R. *et al*. Mars methane detection and variability at Gale crater. *Science (80-.).* **347,** 415–418 (2015).

114. Yamada, T. *et al*. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the. *Int. J. Syst. Evol. Microbiol.* **56,** 1331–1340 (2006).

115. Zahnle, K., Schenk, P.M., Levison, H.F., and Dones, L. (2003) Cratering rates in the outer Solar System. Icarus 163, 263–289.

116. Zahnle K, Haberle RM, Catling DC, Kasting JF (2008) Photochemical instability of the ancient Martian atmosphere. J. Geophys. Res., 113, E11004, doi:10.1029/2008JE003160.

117. Zent AP, McKay CP (1994) The chemical reactivity of the Martian soil and implications for future missions. Icarus, 108, 146–157.

118. Zhang, S. *et al*. Sufficient oxygen for animal respiration 1,400 million years ago. *Proc. Natl. Acad. Sci*. 6pp (2016). doi:10.1073/pnas.1523449113

119. Zolotov, M. Y. & Shock, E. L. Composition and stability of salts on the surface of Europa and their oceanic origin. *J. Geophys. Res. Planets* **106,** 32815–32827 (2001).

Appendix 1

A REDUCED, ABIOTIC NITROGEN CYCLE BEFORE THE RISE OF OXYGEN

Ward, Lewis M, James Hemp, and Woodward W. Fischer. In preparation.

Abstract:

Nitrogen is a critical element for all known life, where it is used in essential biomolecules such as amino and nucleic acids. However, most nitrogen on Earth is found as relatively inert N_2 in the atmosphere, and must be "fixed" to more bioavailable forms before it can be incorporated into biomass. Fixed nitrogen can further be interconverted between a range of oxidized and reduced forms in a complex web of biogeochemical reactions driven by diverse microorganisms in order to conserve energy in addition to using nitrogen for biosynthesis. This redox cycling is a function of the oxidation state of Earth surface environments, and therefore cannot be assumed to have been present on the early Earth, particularly before the evolution of oxygenic photosynthesis and oxygenation of the atmosphere. Here, we consider which steps in the modern nitrogen cycle may have been present on the early Earth before the introduction of molecular oxygen into biology and geochemistry. This includes incorporation of biochemical and phylogenetic evidence for the antiquity of microbial nitrogen metabolisms, as well as geochemical- and photochemical- model based estimates for abiotic processes. We conclude that before the evolution of oxygenic photosynthesis, the nitrogen cycle was largely abiotic, consisting of atmospheric and geological transformations of reduced nitrogen species. Depending on the productivity of the early biosphere, the nitrogen demand of the biosphere may have

been met by abiotic nitrogen fixation processes, consistent with a late origin of biological nitrogen fixation via the nitrogenase enzyme, or an early evolution of nitrogenase for cyanide uptake followed by later cooption for N_2 fixation.

Introduction

The evolution of oxygenic photosynthesis led to the Great Oxygenation Event (GOE) ~2.3 billion years ago (Gya), when O_2 first accumulated to significant concentrations in the atmosphere (Fischer et al. 2016). At this critical juncture, O_2 was introduced into a primarily anoxic global biosphere and geochemical landscape, which in turn revolutionized the nature of the ecology and Earth surface environments as a whole (Ward et al. 2016, Fischer et al. 2016). Before the GOE, oxygenic photosynthesis (and the ability to use water as an electron donor for C fixation) was unavailable to drive high rates of primary productivity that we see today, and O_2 was unavailable for the efficient and complex aerobic heterotrophy and biogeochemical cycles that dominate modern environments (Ward et al. 2017). The size of the biosphere and the structure of biogeochemical cycles before the introduction of O_2 is not well understood, but must have been very different than the modern biosphere fueled by oxygenic photosynthesis and aerobic respiration.

In particular, the pre-GOE, Archean nitrogen cycle must have worked very differently than it does today. Nitrogen is a critical element for all known life, where it is used in essential biomolecules such as amino and nucleic acids. However, most nitrogen on Earth is found as relatively inert N_2 in the atmosphere, and must be “fixed” to more

bioavailable forms before it can be incorporated into biomass (Canfield et al. 2010). Fixed nitrogen can further be interconverted between a range of oxidized and reduced forms in a complex web of biogeochemical reactions driven by diverse microbes that utilize these transformations to conserve energy (Figure 1) in addition to using nitrogen for biosynthesis (Canfield et al. 2010). These reactions include processes that return fixed nitrogen to the atmospheric N_2 pool (denitrification and anammox) making this a closed cycle, maintaining a balance of atmospheric N_2 and fixed N over long timescales (Canfield et al. 2010). This redox cycling, which is a function of the oxidation state of Earth surface environments, cannot be assumed to have been present in the Archean (Ward et al. 2017, Zerkle et al. 2017). For instance, a critical step in the nitrogen cycle—nitrification—has an absolute requirement for O_2 (Klotz and Stein 2008). Nitrification is the only known biological process to oxidize NH_3 to high-valent oxidized N species such as NO_2^- and NO_3^-, which can serve as the substrates for denitrification and anammox to drive the return of fixed nitrogen to N_2 (Mancinelli and McKay 1988). As a result, without O_2 the N cycle is not a closed loop, but instead a vector of nitrogen toward reduced forms.

Furthermore, it is unclear what processes supplied fixed nitrogen to the early biosphere. During recent Earth history, the vast majority of nitrogen fixation has been through biological nitrogen fixation using the nitrogenase enzyme (Canfield et al. 2010); however, it is unclear when nitrogenase evolved. Hypotheses for the origin of nitrogenase range from its presence in the Last Universal Common Ancestor (Fani et al. 2000, Weiss et al. 2016), to evolving much later, close to the GOE (Raymond et al. 2004, Boyd et al.

2011). Without a good understanding of the evolutionary history of biological nitrogen fixation, it is unclear when this process became available to supply fixed nitrogen to the biosphere. Before the evolution of nitrogenase, nitrogen must have been sourced through abiotic mechanisms of converting N_2 into more bioavailable forms, including fixation of N_2 into NO by lightning (Navarro-Gonzalez et al. 1998) or HCN by photochemistry (Tian et al. 2011). The products and fluxes of nitrogen fixation by these abiotic processes is not well constrained, as it is dependent on poorly understood aspects of atmospheric chemistry (such as the ratio of CO_2 to CH_4) (Navarro-Gonzalez et al. 2001). Moreover, because these abiotic fluxes could not be tuned to keep pace with primary productivity as biological nitrogen fixation can (Tyrrell 1999), fixed nitrogen may have eventually become limiting to productivity (Navarro-Gonzalez et al. 2001) or may have proceeded in excess of biological demands and led to depletion of the atmospheric N_2 reservoir (Som et al. 2016) or accumulation of reduced nitrogen in the oceans (Ward et al. 2017). In order to determine when the modern, aerobically driven, biologically closed nitrogen cycle evolved, and what the nature of the earlier nitrogen cycle may have been, we must better constrain the fluxes of nitrogen through the early Earth and the evolutionary history of individual steps in the nitrogen cycle that exist today.

Here, we integrate what is known about the evolution and early history of the nitrogen cycle, and attempt to constrain the drivers and fluxes included in the nitrogen cycle before the introduction of O_2 to biogeochemical cycles following the evolution of oxygenic photosynthesis. We conclude that the modern, aerobic, biologically driven nitrogen cycle is a relatively recent innovation, likely postdating the evolution of

oxygenic photosynthesis and not arising sooner than the Great Oxygenation Event ~2.3 Gya. Before this time, the nitrogen cycle was dominated by reduced species and abiotic processes, primarily the fixation of N_2 to reduced forms by lightning and photochemistry. Reduced nitrogen likely accumulated in the oceans, providing a nutrient source in excess of demands by early electron donor-limited primary productivity. The nitrogen cycle was likely closed by photolysis of a tenuous atmospheric ammonia pool in equilibrium with the ammonia-rich ocean.

Evolution of the biological nitrogen cycle

The modern nitrogen cycle has an absolute requirement for molecular oxygen. Once N_2 is fixed to reduced forms like ammonia, the only known biological pathways for returning nitrogen to the atmosphere are denitrification and anammox, both of which utilize oxidized nitrogen species—NO, nitrite, or nitrate—as substrates. Aerobic nitrification is the only known biological pathway to convert reduced nitrogen into these forms, and so in the absence of molecular oxygen the closed nitrogen cycle collapses to a vector converting N_2 to ammonia.

Nitrification, the sequential oxidation of ammonia to nitrite and nitrate using O_2, is a thermodynamically favorable reaction catalyzed by various groups of microbes. This includes ammonia oxidation by various Proteobacteria and Thaumarchaea, nitrite oxidation by Nitrospirae, Nitrospinae, various Proteobacteria, and one strain of Chloroflexi (Stein and Klotz 2016, Bock and Wagner 2006, Sorokin et al. 2012), and “comammox”, or complete nitrification of ammonia to nitrate, by some Nitrospirae (Daims et al. 2015, van Kessel et

al. 2015). The phylogenetic distribution of these organisms is sparse, and restricted to fairly derived clades, consistent with a relatively late origin and distribution via horizontal gene transfer. Interpretations of a late origin for nitrification are supported by its oxygen requirements, which suggest an origin after the evolution of oxygenic photosynthesis by Cyanobacteria and oxygenation of the atmosphere ~2.3 Gya. Although nitrification has been shown to proceed at oxygen concentrations down to ~10 nM, it exhibits strong oxygen concentration dependence at these ranges, with a half saturation constant of 0.5-1 μM O_2 (Bristow et al. 2016). As a result, nitrification is almost certainly unviable under the trace oxygen concentrations permissible by Archean pO2 proxies (e.g. Johnson et al. 2014). Nitrification, and a complete aerobic nitrogen cycle, may have evolved hand-in-hand with the GOE (e.g. Zerkle et al. 2017), but not before.

In the absence of nitrification, what then would be the topology of the early biological nitrogen cycle? Were alternative mechanisms in place to produce oxidized nitrogen, or was the nitrogen cycle restricted to reduced forms? These questions can be addressed by combining insights from the rock record about abiotic transformations with comparative biology analyses of the relative age of nitrogen metabolisms. A first order question that must first be answered is whether in the absence of aerobic nitrification as we see it today other sources, either abiotic or biological, could supply oxidized nitrogen species to the early biosphere.

It has been proposed that atmospheric processes, particularly lightning, could fix N_2 into bioavailable forms on the early Earth (Nacarro-Gonzalez et al. 1998). Depending on atmospheric chemistry, the products of this abiotic fixation could range from oxidized

forms like NO (under high CO_2:CH_4 ratios) to reduced phases like HCN (under relatively high methane atmospheres) (Navarro-Gonzalez et al. 2001). The fluxes of nitrogen fixed by these processes could be significant for much of Earth history (Ward et al. 2017), and oxidized nitrogen produced through these processes has been proposed as the source of the first high potential electron acceptors to the biosphere (Ducluzeau et al. 2009). The high pCO_2 necessary for these sources to produce significant oxidized nitrogen, however, are inconsistent with several proxies (e.g. Rye et al. 1995, Blättler et al. 2016), while numerous models suggest that methane was a more significant greenhouse gas in counteracting the Faint Young Sun (e.g. Pavlov et al. 2000, Kasting et al. 2001, Kasting 2005); as a result, abiotic supplies of oxidized nitrogen at this time cannot be assumed to be significant. Even if abiotic fluxes of oxidized nitrogen existed on the early Earth, it is possible that the compounds produced would not have been stable over long enough timescales to be useful to the early biosphere. The Archean ocean is thought to have been relatively rich in dissolved ferrous iron (Holland 1984), and many oxidized nitrogen species, including nitrite and NO, react rapidly with ferrous iron (Klueglein and Kappler 2012, Kopf et al. 2013). As a result, any oxidized nitrogen reaching the Archean ocean from atmospheric processes may have been reduced on timescales of minutes, preventing it from playing a role in the biological nitrogen cycle until after the Rise of Oxygen titrated reduced iron from the oceans.

An alternative route to oxidized nitrogen species in the Archean ocean are uncharacterized anaerobic nitrification pathways. These include the oxidation of ammonia coupled to the reduction of iron oxides (feammox) or to phototrophy (photoammox). Many

novel nitrogen metabolisms have been uncovered in recent years, including aerobic methane oxidation coupled to denitrification (Kits et al. 2015, Skennerton et al. 2015), intraaerobic methane oxidation driven by nitric oxide dismutation (Ettwig et al. 2012), and the combined oxidation of ammonia to nitrite and then nitrate (comammox, van Kessel et al. 2015, Daims et al 2015). It is therefore reasonable that additional nitrogen metabolisms exist in the environment and simply have not yet been discovered. In 1977, Broda predicted the existence of anammox on the basis of thermodynamic calculcations, proving to be one of the most prescient hypotheses in environmental microbiology. In the same paper, he also proposed the existence of photoammox, yet this metabolism has still not been identified in an organism. In order to oxidize ammonia to nitrite or N_2, a total of 6 electrons would need to be transferred without the release of toxic intermediates like hydroxylamine; this is unheard of in the distribution of characterized phototrophs, and so may not be possible. While the phototrophic oxidation of nitrite to nitrate has been described in a small number of Proteobacteria (Griffin et al. 2007, Schott et al. 2010), this metabolism has been shown to have evolved relatively recently (Hemp et al. 2016) and therefore does not reflect an early history of phototrophic nitrification. Further investigation may uncover photoammox organisms in modern or ancient environments, but as of now there is no evidence to suggest that this metabolism has ever evolved.

Similarly, the oxidation of ammonia coupled to iron oxide reduction (feammox) has been proposed but never demonstrated in an isolated organism. This metabolism is thermodynamically favorable under a range of conditions, and would be a logical component of an Archean ecosystem driven by the phototrophic oxidation of iron (Fischer

and Knoll 2009). Evidence for the occurrence of feammox has been described in environments such as wetlands (Clement et al. 2005, Yang et al. 2012), but no conclusive evidence has been presented, nor has a microbe capable of driving this process yet been described. This metabolism, while thermodynamically favorable, is significantly kinetically inhibited. In ammonia oxidizing organisms, ammonia monooxygenase activates ammonia to more accessible forms in an O_2-dependent reaction (Bock and Wagner 2006). The process of methane oxidation is evolutionarily related to ammonia oxidation, using homologous methane monooxygenase enzymes. While some methane oxidizing bacteria are capable of using alternative electron acceptors during methanotrophy (e.g. Kits et al. 2015, Skennerton et al. 2015), these strains retain an absolute requirement for trace O_2 for the initial activation of methane, suggesting that this step is challenging if not impossible to achieve in a fully anaerobic world. The most likely scenario for an anaerobic nitrifier is one analogous to the anaerobic oxidation of methane, achieved by specialized archaea running a process of reverse methanogenesis that is still not fully understood (McGlynn 2017). It is conceivable that a pathway of reverse Dissimilatory Nitrate Reduction to Ammonia could exist, though it has never been demonstrated. Any mechanism of ammonia oxidation without O_2 or oxidized nitrogen as a substrate would therefore require exotic biochemistry, acting in an unknown organism.

It is thermodynamically favorable for feammox to proceed without biological catalysis; however, iron oxides may not have been abundant in the Archean ocean (Ward et al. 2017), and the reaction of iron oxides with ammonia is kinetically inhibited and does not

proceed spontaneously under reasonable timescales under ocean-relevant conditions (Supplemental Information).

Hypotheses about potential sources of oxidized nitrogen on the early Earth can be tested by querying the biological record of metabolisms that make use of these compounds. If a source of oxidized nitrogen was present before the Rise of Oxygen, then metabolisms that make use of these compounds, such as denitrification, should appear to be evolutionarily ancient, predating the evolution of aerobic respiration. If, however, nitrogen respiration appears to have evolved post-O_2, this will increase our confidence that no significant oxidized nitrogen was available to the early biosphere.

In order for the nitrogen cycle to be maintained at steady state over geological timescales, fixation of nitrogen must be balanced by return of nitrogen to the atmospheric N_2 pool. Today, this return is accomplished via denitrification and anammox. These metabolisms make use of oxidized nitrogen species (e.g. nitrite and nitrate) as respiratory electron acceptors coupled to the oxidation of organic carbon or ammonia.

Denitrification is the process of respiratory reduction of nitrate to N_2 via nitrite, nitric oxide, and nitrous oxide. This is a multistep process that can be found as either complete or partial pathways in a range of microbial taxa (Klotz and Stein 2016), utilizing a variety of enzymes for donating electrons from the electron transport chain onto nitrogen. The distribution of denitrification genes appears to be the result of extensive horizontal gene transfer, of entire pathways and individual genes (Jones et al. 2008). Despite the thermodynamic favorability of oxidized nitrogen species as electron acceptors, not all steps in denitrification are coupled to energy conservation, making it a relatively inefficient

pathway in terms of protons translocated per electron (Chen and Strous 2012). Some of the steps in the denitrification pathway are catalyzed by only a single enzyme or by a small group of evolutionarily related proteins. This includes nitric oxide reduction, which is catalyzed by several families of closely related Heme Copper Oxidoreductase enzymes (Hemp et al., in prep). The HCO superfamily is primarily made up of O_2 reductases, but several independent lineages of these enzymes have convergently evolved the ability to catalyze nitrogen chemistry, including putative nitric oxide dismutation, nitrous oxide reduction, and most often nitric oxide reduction (Figure 2)(Hemp et al. in prep). Evolutionary analysis of the origins of NOR enzymes can help reveal the antiquity of denitrification, and, by extension, the oxidized half of the nitrogen cycle.

Structural and phylogenetic analysis of the HCO superfamily has revealed that the A family O_2 reductases, used for aerobic respiration at relatively high oxygen concentrations (>1 micromolar), with nitric oxide reductases evolving from the derived B and C family family O_2 reductases which appear to be secondarily adapted to low oxygen concentrations, having exchanged a conserved proton channel for an oxygen channel, increasing the diffusion of O_2 to the active site at the cost of protons pumped (Hemp et al., in prep).

Nitrogen fixation on the early Earth

The foundational step in the nitrogen cycle is the fixation of atmospheric N_2 to reduced forms that can be incorporated into biomass. On the modern Earth, this is accomplished biologically via the nitrogenase enzyme (Canfield et al. 2010). It is unclear, however,

exactly when biological nitrogen fixation evolved, and therefore whether the biosphere relied on abiotically fixed nitrogen for some period of its early history. The evolution of biological nitrogen fixation via nitrogenase is contested, with estimates for its origin spanning nearly 2 billion years (Fani et al. 2000, Raymond et al. 2004, Boyd et al. 2011, Suteken et al. 2015, Boyd et al. 2015, Weiss et al. 2016). On the early Earth, before the invention of oxygenic photosynthesis, global rates of primary productivity were likely limited by geological fluxes of electron donor compounds like H_2 (Ward et al. 2017). As a result, the efficiency of consumption of these electron donors would be a critical constraint on the overall rates of productivity that result. Nitrogen fixation is an energetically costly process, and for the energy- and electron-limited metabolisms that fueled the Archean biosphere, this process would have likely been a major drain on productivity. It is therefore a reasonable expectation that nitrogenase might not evolve without strong selective pressure for biological nitrogen fixation, likely as a result of the nitrogen demand of biological primary productivity outpacing nitrogen supply through abiotic processes. To constrain when this is likely to occur, we must integrate expectations of abiotic nitrogen fixation processes compared against estimates of primary productivity through time. Abiotic nitrogen fixation on the early Earth likely occurred through a variety of processes, including lightning, photochemistry, and hydrothermal reactions (Ward et al. 2017). While the fluxes through each of these process is not well constrained, and relies on poorly understood aspects of the early Earth system like the CO_2:CH_4 ratio of the atmosphere (Navarro-Gonzalez et al. 2001), their sum can be estimated to within about an order of

magnitude, which is sufficient accuracy for comparison with expectations of early biological productivity.

It has been hypothesized that nitrogenase was present in the last universal common ancestor of life on Earth (Raymond et al 2004, Weiss et al. 2016), but this interpretation is widely disputed (Boyd et al 2011). For much of the Archean Era, abiotic nitrogen fixation through lightning and other processes could have supplied a sufficient nitrogen flux to support the Earth's much less productive biosphere (Vlaeminck et al 2011, Kharecha et al 2005, Ward et al. 2017), but increases in primary productivity and potentially declines in abiotic nitrogen fixation fluxes in the latest Archean and early Paleoproterozoic may have led to a decrease in abiotic nitrogen fixation, possibly driving the evolution of nitrogenase (Navaro-Gonzalez et al 2001, Ward et al. 2017). This nitrogen crisis may have been exacerbated by the Great Oxidation Event, as the rise in atmospheric O_2 would have allowed the nitrification of a large marine ammonium pool that may have developed in the Archean, followed by denitrification and massive loss of accumulated fixed nitrogen (Fennel et al 2005, Ward et al. 2017). Nitrogenase may then have evolved around the GOE (Boyd et al 2011). An alternative scenario for the evolution of nitrogenase is that it initially evolved for the uptake of HCN, produced via atmospheric processes under a high CH_4 atmosphere (Ward et al. 2017). Nitrogenase can be used for the uptake of HCN (Materassi and Balloni 1977, Li et al. 1982, Dekas et al. 2009), and so large fluxes of HCN to the oceans may have led to the evolution of nitrogenase first as a way to detoxify and take up HCN, and was only later coopted to N_2 fixation (Silver and Postgate 1973, Raymond 2005, Ward et al. 2017).

HCN produced via atmospheric processes (e.g. lightning and photochemistry) is expected to be delivered to the oceans by rain on timescales of ~10 years (Zahnle 1986). HCN could then be converted to ammonia via hydrolysis over timescales of decades (Abelson 1966) or directly taken up by biology via nitrogenase (Raymond 2005). Depending on pH, dissolved HCN could even polymerize and subsequently hydrolyze to produce glycine or other amino acids (Abelson 1966). Availability of atmospherically fixed nitrogen to the biosphere would depend partially on the precipitation of cyanide as ferrocyanide ($Fe(CN)_6)^{4-}$) upon reaction with dissolved iron; the production of this compound would depend on the concentration of dissolved iron in the surface ocean, which may have been low relative to the average ocean at this time (Ward et al. 2017) and relative timescales of mixing and cyanide hydrolysis. It is not known if nitrogen in ferrocyanide is accessibly to biology, and so the reaction dynamics and uptake of this phase via nitrogenase could be an important constraint for understanding the supply of nitrogen to the early biosphere. If hydrolysis is relatively slow, this would potentially encourage the early evolution of nitrogenase for the uptake of nitrogen from cyanide or ferrocyanide. However, if conversion of cyanide to ammonia, and subsequent ammonia recycling, is efficient, then the nitrogen demands of the biosphere may have been met or exceeded without nitrogenase.

An abiotic nitrogen cycle before oxygenic photosynthesis

Based on the above discussion, it appears that most steps in the biological nitrogen cycle were not viable before the evolution of oxygenic photosynthesis. Nitrification,

denitrification, and anammox were not present, and biological nitrogen fixation may or may not have evolved. However, the biosphere still needed fixed nitrogen, and so for the biosphere and an N_2-rich atmosphere to coexist a dynamic steady state must have existed between reduced nitrogen species accessible to biology and return of N_2 to the atmosphere. This could have taken the form of a largely abiotic, reduced nitrogen cycle before the evolution of oxygenic photosynthesis. In this model (Figure 3), atmospheric N_2 is fixed to more reduced forms by abiotic processes including lightning, photochemistry, and reactions at hydrothermal vents; organic nitrogen is interconverted with biomass; small fluxes of organic nitrogen are buried and eventually returned to the atmosphere as N_2 following subduction and degassing, but most nitrogen is returned to the atmosphere via equilibration of marine ammonia with the atmosphere and subsequent photolysis. The primary pools of nitrogen would therefore be atmospheric N_2, marine NH_3, organic nitrogen, and atmospheric NH_3. Assuming previous estimates of abiotic nitrogen fixation rates (Ward et al. 2017), burial of organic nitrogen estimated by a 40 C:1 N ratio (Berner 2006) scaled to expected organic carbon burial (Ward et al. 2017), volcanic outgassing rates of nitrogen (Sano et al. 2001), equilibration between dissolved and gaseous ammonia, and estimates of photolysis of atmospheric ammonia (Kasting 1982), a simple steady state model of the Archean nitrogen cycle can be developed. Under these steady state conditions, fixation of N_2 to reduced forms results in accumulation of NH_3 in the oceans, likely more than meeting the nitrogen demands of the electron donor-limited early biosphere, even without recycling or biological nitrogen fixation (Ward et al. 2017). This dissolved ammonia will be in equilibrium with a tenuous atmospheric ammonia phase, which would

be rapidly photolysed back to N_2. Depending on atmospheric composition, temperature, and other variables, dissolved ammonia concentrations in the ocean would range over the order of 1-100 μM, with atmospheric ammonia mixing ratios on the order of 10^{-8}—too low for ammonia to play a significant role in maintaining climate under the faint young sun (Kasting 1982). If significant haze developed in the Archean atmosphere, NH_3 may have been shielded from UV photolysis, increasing the proportion of nitrogen that could be locked into reduced species (Pavlov et al. 2001).

Conclusions

Under the scenario described above, the early, pre-oxygen nitrogen cycle was largely abiotic, consisting of transformations between N_2 and reduced nitrogen species primarily via atmospheric processes. Life's role would have been minor, and consisted largely of interconversion of reduced nitrogen with biomass, and delivery of a portion of that organic nitrogen to sediments. Given the expected low productivity of the early biosphere, the relatively high expected abiotic nitrogen fixation rates, and the likelihood of high nitrogen recycling, biological nitrogen fixation may not have been necessary until relatively late in Earth history, potentially evolving around the time of the Great Oxygenation Event. Alternatively, nitrogenase may have evolved earlier, potentially in a niche nitrogen-limited environment or for the uptake of cyanide into biomass followed later by cooption for N_2 fixation, a scenario that has been hypothesized previously (e.g. Silver and Postgate 1973, Raymond 2005, Ward et al. 2017). Whether nitrogenase predated the GOE or not, it appears that the remainder of the nitrogen cycle did not evolve until this

time. Tests of this hypothesis will hinge on development of an appropriate isotope mass balance framework for interpreting the Archean nitrogen cycle in the context of the fluxes described here.

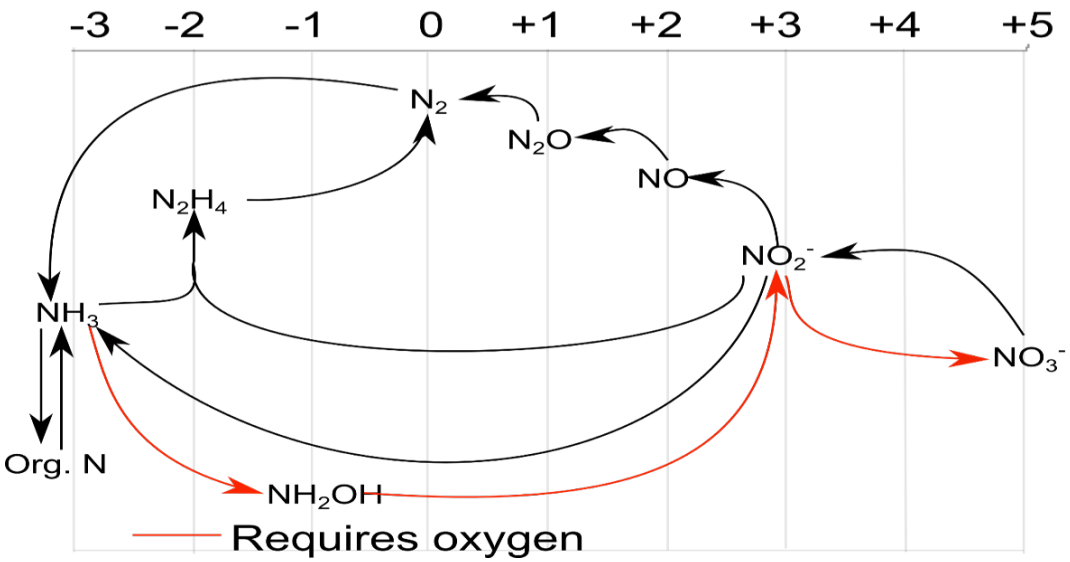


Figure 1: Topology of the nitrogen cycle plotted against the oxidation state of nitrogen. Atmospheric N2 can be fixed into reduced forms like ammonia that can be interconverted with organic nitrogen. Subsequent cycling requires molecular oxygen to drive nitrification to nitrite and nitrate. These compounds can then serve as substrates for denitrification and anammox, returning nitrogen to the atmosphere. Without O2, however, nitrogen would become trapped in reduced forms.

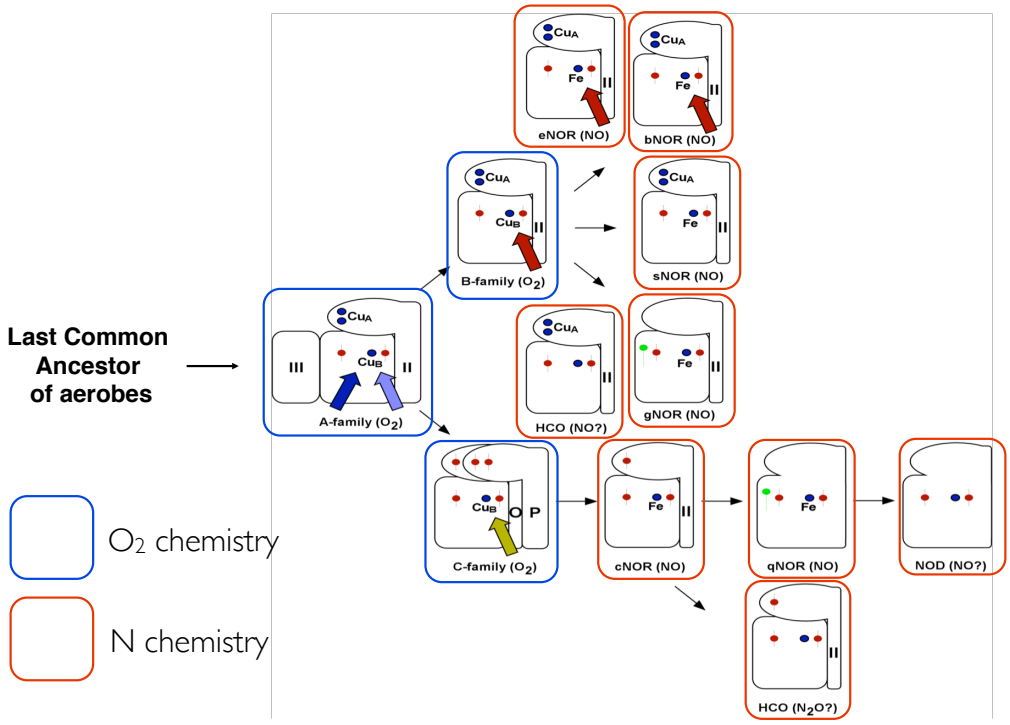


Figure 2: Cartoon of the evolution of the Heme Copper Oxidoreductase superfamily, showing derivation of nitrogen reductases from the A-family O2 reductases. Modified from Hemp et al.

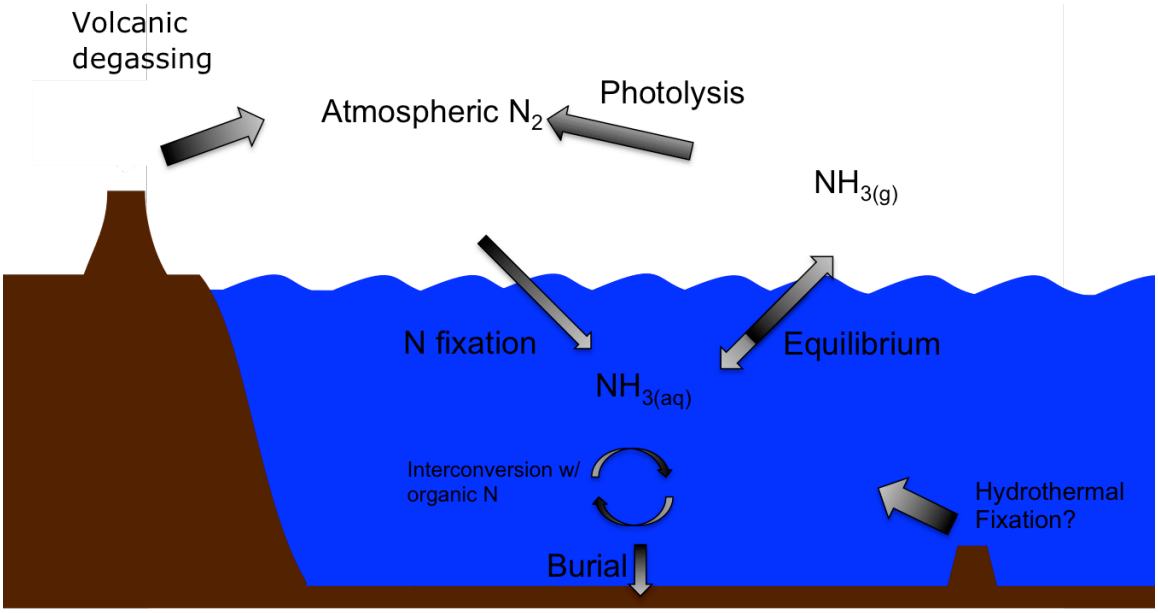


Figure 3: Topology of the pre-oxygen, Archean nitrogen cycle described above. Nitrogen existed in only reduced forms, and was primarily cycled between NH_3 and N_2 via abiotic atmospheric processes.

References

1. Abelson, P., Chemical events on the primitive Earth, Proc. Nat..,icad. Sci. USA, 55, 1365-1372, 1966
2. Berner, R. A. Geological nitrogen cycle and atmospheric N2 over Phanerozoic time. Geology 34, 413–414 (2006).
3. Blättler, C. et al. Nat. Geosci. 1, (2016).
4. Bock, Eberhard, and Michael Wagner. "Oxidation of inorganic nitrogen compounds as an energy source." The prokaryotes. Springer New York, 2006. 457-495.
5. Boyd, ES et al. 2011. “A late methanogen origin for molybdenum-dependent nitrogenase”. Geobiology 9 pp221-232.
6. Boyd, ES et al. 2011b. “An alternative path for the evolution of biological nitrogen fixation”. Frontiers in Microbiology 2(205).
7. Boyd, E. S., Garcia Costas, A. M., Hamilton, T. L., Mus, F. & Peters, J. W. Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. J. Bacteriol. 197, JB.02611-14 (2015).
8. Bristow, Laura A., et al. "Ammonium and nitrite oxidation at nanomolar oxygen concentrations in oxygen minimum zone waters." Proceedings of the National Academy of Sciences 113.38 (2016): 10601-10606.
9. Broda, E. "Two kinds of lithotrophs missing in nature." Zeitschrift für allgemeine Mikrobiologie 17.6 (1977): 491-493.

10. Canfield, D. E., Glazer, A. N. & Falkowski, P. G. The evolution and future of Earth's nitrogen cycle. Science 330, 192–196 (2010).

11. Chen, J and M Strous. 2012. "Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution". Biochimica et Biophysica Acta (BBA)-Bioenergetics 1827(2) pp136-144.

12. CLEMENT, J., SHRESTHA, J., EHRENFELD, J. & JAFFE, P. Ammonium oxidation coupled to dissimilatory reduction of iron under anaerobic conditions in wetland soils. Soil Biol. Biochem. 37, 2323–2328 (2005).

13. Daims, H. et al. Complete nitrification by Nitrospira bacteria. Nature 528, 504–509 (2015).

14. Ducluzeau, A. L. et al. Was nitric oxide the first deep electron sink? Trends Biochem. Sci. 34, 9–15 (2009).

15. Ettwig, K. F. et al. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature 464, 543–548 (2010).

16. Fani, Renato, Romina Gallo, and Pietro Lio. "Molecular evolution of nitrogen fixation: the evolutionary history of the nifD, nifK, nifE, and nifN genes." *Journal of Molecular Evolution* 51.1 (2000): 1-11.

17. Fennel, K et al. 2005. "The co-evolution of the nitrogen, carbon and oxygen cycles in the Proterozoic ocean". American Journal of Science 305 pp526-545.

18. Fischer, W. W. & Knoll, A. H. An iron shuttle for deepwater silica in late Archean and early Paleoproterozoic iron formation. Bull. Geol. Soc. Am. 121, 222–235 (2009).

19. Fischer, W. W., Hemp, J. & Johnson, J. E. Manganese and the Evolution of Photosynthesis. Orig. Life Evol. Biosph. (2015). doi:10.1007/s11084-015-9442-5

20. Griffin, B. M., Schott, J. & Schink, B. Nitrite, an electron donor for anoxygenic photosynthesis. Science 316, 1870 (2007).

21. Hemp, J. et al. Genomics of a phototrophic nitrite oxidizer: insights into the evolution of photosynthesis and nitrification. ISME J. 1–10 (2016). doi:10.1038/ismej.2016.56

22. Hemp, J, R Murali, LA Pace, RA Sanford, LM Ward, WW Fischer, and RB Gennis. Diversity and evolution of nitric oxide reduction. In prep.

23. Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press

24. Johnson, J. E., Gerpheide, A., Lamb, M. P. & Fischer, W. W. O2 constraints from Paleoproterozoic detrital pyrite and uraninite. Geol. Soc. Am. Bull. 126, 813–830 (2014).

25. Jones, CM et al. 2008. “Phylogenetic analysis of nitrite, nitric oxide, and nitrous oxide respiratory enzymes reveal a complex evolutionary history for denitrification”. Mol. Biol. Evol. 25(9) pp1955-1966.

26. Kasting, J. F. Stability of ammonia in the primitive terrestrial atmosphere. J. Geophys. Res. 87, 3091–3098 (1982).

27. Kasting, J. F., Pavlov, A. A. & Siefert, J. L. A coupled ecosystem-climate model for predicting the methane concentration in the Archean atmosphere. Orig. Life Evol. Biosph. 31, 271–285 (2001)

28. Kasting, J. F. Methane and climate during the Precambrian era. Precambrian Res. 137, 119–129 (2005).

29. Kharecha, P., Kasting, J. & Siefert, J. A coupled atmosphere–ecosystem model of the early Archean Earth. Geobiology 3, 53–76 (2005).

30. Kits, K. D., Klotz, M. G. & Stein, L. Y. Methane oxidation coupled to nitrate reduction under hypoxia by the Gammaproteobacterium Methylomonas denitrificans, sp. nov. type strain FJG1. Environ. Microbiol. n/a–n/a (2015).

31. Li, Jiage, Barbara K. Burgess, and James L. Corbin. "Nitrogenase reactivity: cyanide as substrate and inhibitor." *Biochemistry* 21.18 (1982): 4393-4402.

32. Mancinelli, R. L. & McKay, C. P. The evolution of nitrogen cycling. Origins of Life and Evolution of the Biosphere 18, 311-325 (1988).

33. Materassi, R., and W. Balloni. "Cyanide Reduction by Nitrogenase in Intact Cells of Rhodopseudomonas gelatinosa Molisch." *Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene. Zweite Naturwissenschaftliche Abteilung: Allgemeine, Landwirtschaftliche und Technische Mikrobiologie* 132.5-6 (1977): 413-417.

34. McGlynn, Shawn E. "Energy Metabolism during Anaerobic Methane Oxidation in ANME Archaea." Microbes and Environments 32.1 (2017): 5-13.

35. Navaro-Gonzalez, R et al. 2001. “A possible nitrogen crisis for Archean life due to reduced nitrogen fixation by lightning”. Nature 412 pp61-64.

36. Navarro-Gonzalez, R et al. Geophys. Res. Lett. 25, 3123 (1998).

37. Pavlov, A. A., Kasting, J. F., Eigenbrode, J. L. & Freeman, K. H. Organic haze in Earth’s early atmosphere: Source of low-13C Late Archean kerogens? Geology 29, 1003–1006 (2001).

38. Pavlov, A., Kasting, F., Brown, L. L., Rages, K. & Freedman, R. Greenhouse warming by CH4 in the atmosphere of early Earth. J. Geophys. Res. 105, 11981–11990 (2000).

39. Raymond, J., Siefert, J. L., Staples, C. R. & Blankenship, R. E. The Natural History of Nitrogen Fixation. Mol. Biol. Evol. 21, 541–554 (2004).

40. Rye, R., Kuo, P. H. & Holland, H. D. Atmospheric carbon dioxide concentrations before 2.2 billion years ago. Nature 378, 603–605 (1995).

41. Sano, Y., Takahata, N. & Nishio, Y. Volcanic flux of nitrogen from the Earth. Chem. Geol. 171, 263–271 (2001).

42. Schott, J., Griffin, B. M. & Schink, B. Anaerobic phototrophic nitrite oxidation by Thiocapsa sp. strain KS1 and Rhodopseudomonas sp. strain LQ17. Microbiology 156, 2428–2437 (2010).

43. Skennerton, C. T. et al. Genomic reconstruction of an uncultured hydrothermal vent gammaproteobacterial methanotroph (family Methylothermaceae) indicates multiple adaptations to oxygen limitation. Front. Microbiol. 6, 1–12 (2015).

44. Som, S. M. et al. Nat. Geosci. 9, 1–5 (2016).

45. Tian, F., Kasting, J. F. & Zahnle, K. Revisiting HCN formation in Earth's early atmosphere. Earth Planet. Sci. Lett. 308, 417–423 (2011).

46. Tyrrell, T. 1999. "The relative influences of nitrogen and phosphorous on oceanic primary production". Nature 400 pp525-531.

47. van Kessel, M. A. H. J. et al. Complete nitrification by a single microorganism. Nature 528, 1–17 (2015).

48. Vlaeminck, SE et al. 2011. "In quest of the nitrogen oxidizing prokaryotes of the early Earth". Environmental Microbiology 13(2) pp283-295.

49. Ward, L. M., Kirschvink, J. L. & Fischer, W. W. Timescales of Oxygenation Following the Evolution of Oxygenic Photosynthesis. Orig. Life Evol. Biosph. 46, 51–65 (2016).

50. Ward, et al. 2017. Electron donor supply limited primary productivity before the evolution of oxygenic photosynthesis. In prep.

51. Weiss, M. C. et al. The physiology and habitat of the last universal common ancestor. Nat. Microbiol. 1–8 (2016).

52. Yang, W. H., Weber, K. A. & Silver, W. L. Nitrogen loss from soil through anaerobic ammonium oxidation coupled to iron reduction (SI). Nat. Geosci. 5, 538–541 (2012).

53. Zahnle, K. J. Photochemistry of methane and the formation of hydrocyanic acid (HCN) in the Earth's early atmosphere. J. Geophys. Res. 91, 2819–2834 (1986).

54. Zerkle, A et al. Onset of the aerobic nitrogen cycle during the Great Oxidation Event. Nature 1–10 (2017).

Supplemental Information:

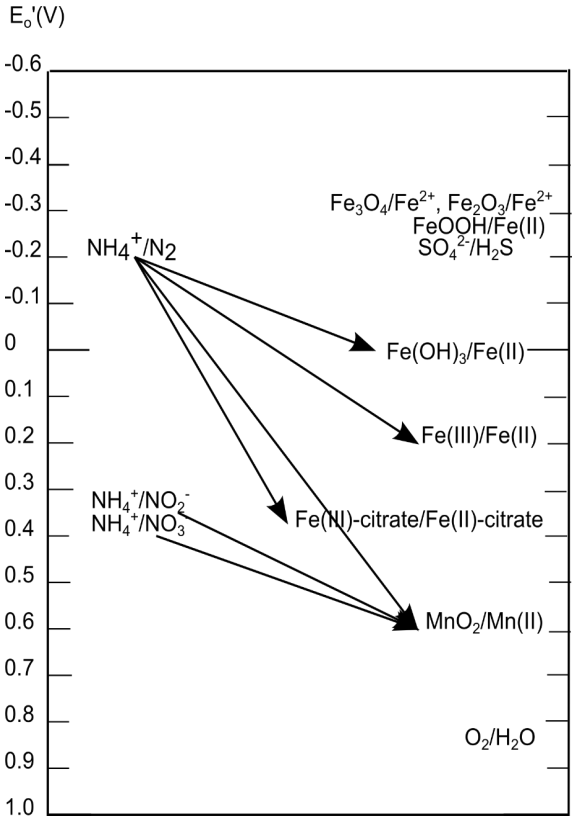
Anaerobic oxidation of ammonium coupled to Fe(III) reduction— "feammox"—could have provided a source of oxidized nitrogen before the evolution of oxygenic photosynthesis, and therefore promoted the closure of the nitrogen cycle without oxygen. Production of N_2 or nitrite by this process is thermodynamically favorable at environmentally reasonable substrate (Supplemental Figure 1, Supplemental Figure 2). While no microorganism has ever been characterized as driving this process, it is conceivable that it could proceed abiotically. To characterize the kinetics of the abiotic reactions, we incubated sterile solutions of ammonium with ferrihydrite (a reactive, environmentally common, poorly-ordered ferric iron phase) and measured the production of reduced iron over time. Results suggest that feammox does not proceed spontaneously

under environmentally relevant conditions, so this process is unlikely to play a role in modern or ancient environments in the absence of biological catalysis. In contrast, similar abiotic incubations containing manganese oxides demonstrate spontaneous reaction with ammonia over timescales of several months. This process ("manoxammox") may be important in some modern environments (Hulth et al. 1999, Luther et al. 1997) but was unlikely to be a factor in the nitrogen cycle before the evolution of oxygenic photosynthesis, as manganese oxides were not present on Earth's surface until around the GOE (Johnson et al. 2013).

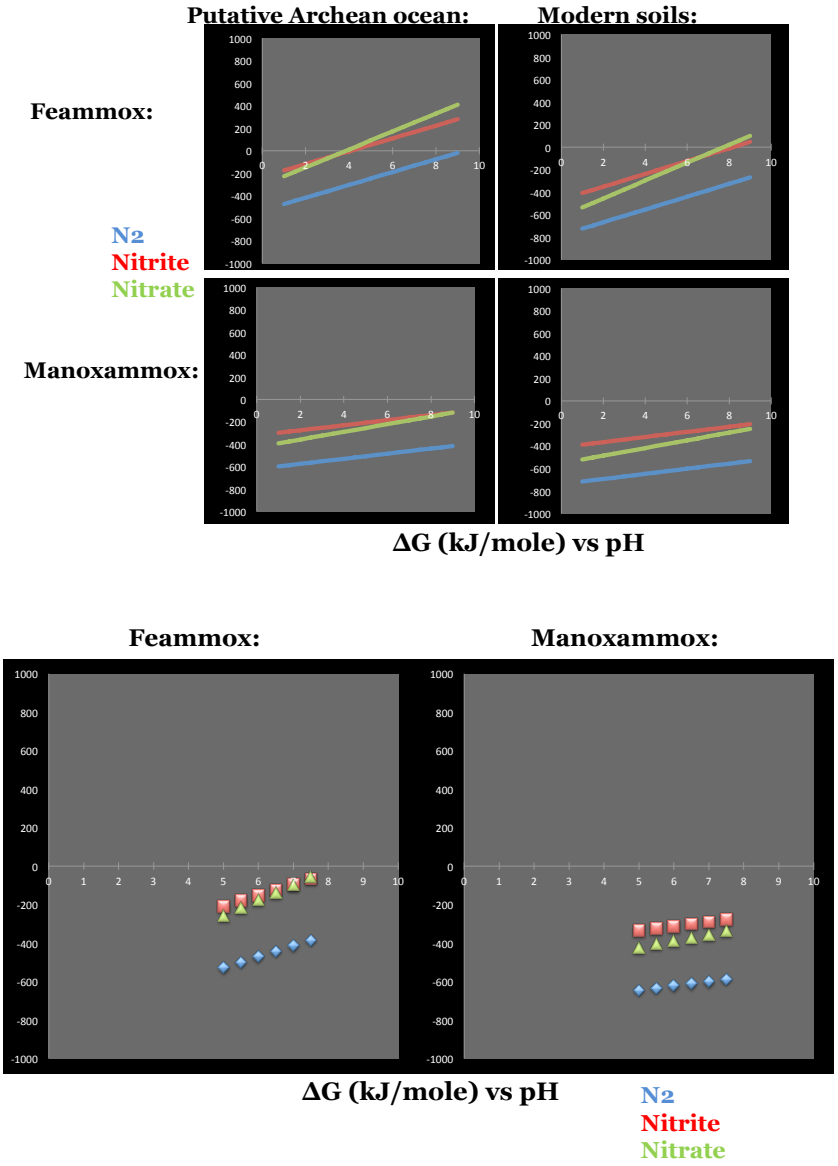
Experiments consisted of serum vials prepared anaerobically with milliQ H_20, buffered with 10mM PIPES and adjusted to pH 5 or 7, to which were added 5-100 mM NH_4Cl and 1 g/L of ferrihydrite (feammox), or either birnessite or colloidal MnO_2 (manoxammox). Vials were then sparged with 100% N_2 and overpressured to 20 psi. Vials were left in the dark at room temperature for the duration of the experiment except during sampling, which was performed in an anaerobic chamber. Measurements were taken via ion chromatography on a Dionex DX500 equipped parallel CS16 cation and AS19 anion columns to measure Mn(II), NO_3^-, NO_2^-, and NH_4^+. Fe(II) was measured using the ferrozine assay (Carter 1971). Sensitivity was linear down to analyte concentrations of ~50μM, and consistent to within ~3%.

No change was detected in the concentrations of any reactants during weekly sampling for 70 days (Supplemental Figure 3). However, following several additional months additional incubation in the dark at room temperature, dissolution of colloidal manganese was determined by visual assessment, likely via reduction by ammonium (Supplemental

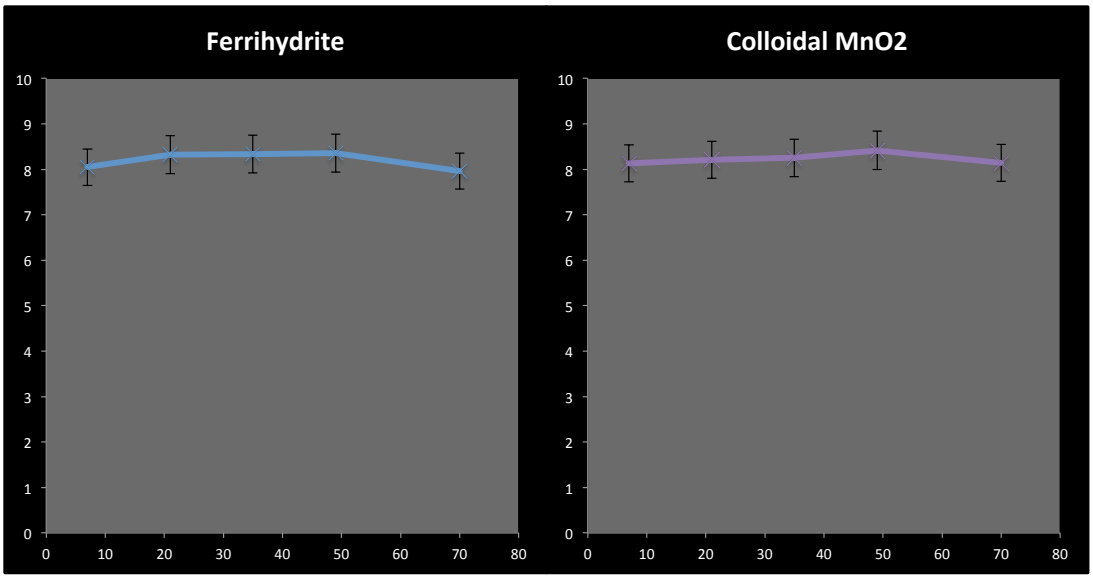
Figure 4). This suggests that manoxammox can proceed spontaneously over timescales of several months, despite no measurable reaction in the first several weeks. No visible change in ferrihydrite or birnessite was observed even up to three years later, indicating that while some poorly ordered manganese oxide phases are capable of spontaneous reaction with ammonia, this is not true of iron oxides or more highly ordered manganese.



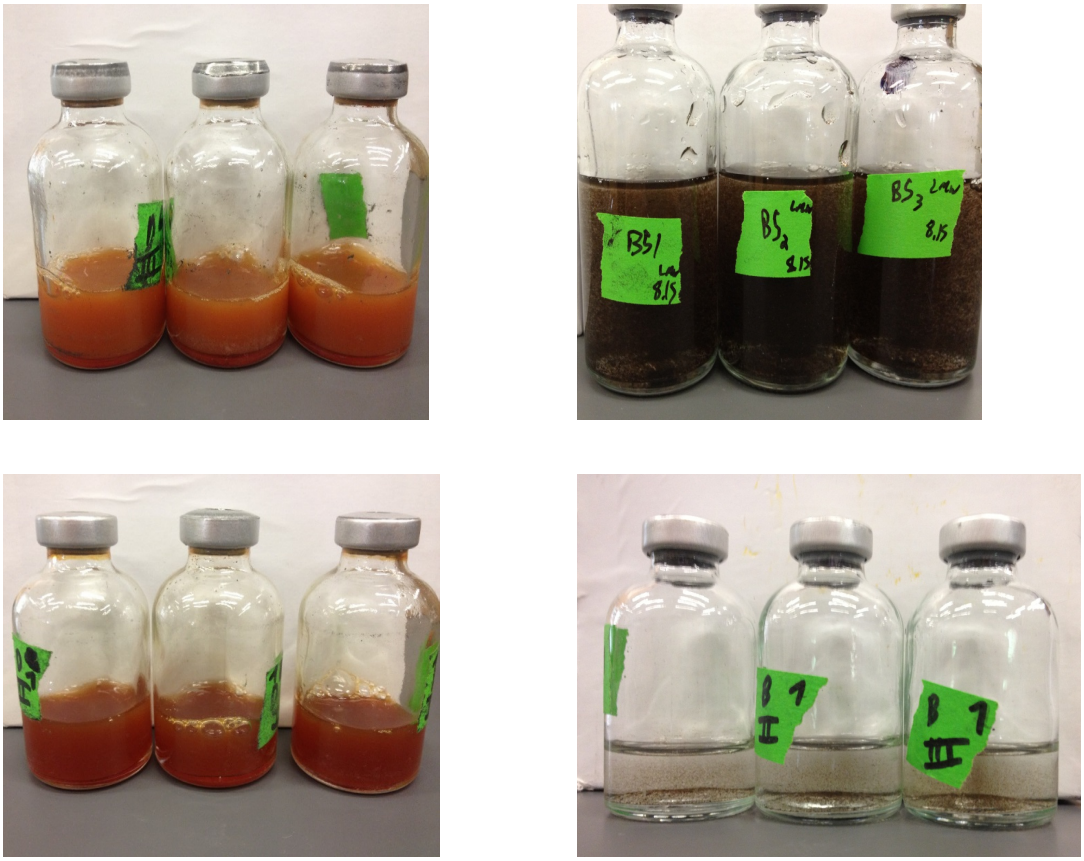
Supplemental Figure 1: Redox ladder with reduction potential of feammox/manoxammox reactions (arrows). Values are calculated at standard state and pH 7.



Supplemental Figure 2: Thermodynamic favorability (ΔG/kJ/mole) against pH for feammox under putative Archean conditions (top left), feammox in modern iron-rich anoxic soils (top right), manoxammox under putative late Archean/early Proterozoic conditions (middle left), and manoxammox in modern manganese-rich anoxic soils (middle right). Below, favorability of feammox and manoxammox, respecitively, under experimental conditions using ferrhydrite and birnessite.



Supplemental Figure 3: Measurements of dissolved ammonium concentrations over 70 days of incubation. No significant change in ammonium was detected in this time.



Supplemental Figure 4: Photographs of incubations. Top, vials as initially prepared, containing dense suspended metal oxides (ferrihydrite at left, colloidal manganese at right). Below, vials after several months of incubation. Ferrihydrite solutions have not noticeably reacted, while colloidal manganese has largely dissolved.

Supplemental references:

1. Carter, P. Spectrophotometric determination of serum iron at the submicrogram level with a new reagent (ferrozine). *Anal. Biochem.* **40,** 450–458 (1971).
2. Hulth, S., Aller, R. C. & Gilbert, F. Coupled anoxic nitrification manganese reduction in marine sediments. *Geochim. Cosmochim. Acta* **63,** 49–66 (1999).
3. Luther, G. W., Sundby, B., Lewis, B. L., Brendel, P. J. & Silverberg, N. Interactions of manganese with the nitrogen cycle: Alternative pathways to dinitrogen. *Geochim. Cosmochim. Acta* **61,** 4043–4052 (1997).

Appendix 2

CONE-FORMING CHLOROFLEXI MATS AS ANALOGS OF CONICAL STROMATOLITE FORMATION WITHOUT CYANOBACTERIA

Lewis M. Ward, Woodward W. Fischer, Katsumi Matsuura, and Shawn E. McGlynn. In preparation.

Abstract

Modern microbial mats provide useful process analogs for understanding the mechanics behind the production of ancient stromatolites. However, studies to date have focused on mats composed predominantly of oxygenic Cyanobacteria (Oxyphotobacteria) and algae, which makes it difficult to assess a unique role of oxygenic photosynthesis in stromatolite morphogenesis, versus different mechanics such as phototaxis and filamentous growth.

Here, we characterize Chloroflexi-rich hot spring microbial mats from Nakabusa Onsen, Nagano Prefecture, Japan. This spring supports cone-forming microbial mats in both upstream high-temperature, sulfidic regions dominated by filamentous anoxygenic phototrophic Chloroflexi, as well as downstream Cyanobacteria-dominated mats. These mats produce similar morphologies analogous to conical stromatolites despite metabolically and taxonomically divergent microbial communities as revealed by 16S and shotgun metagenomic sequencing and microscopy. These data illustrate that anoxygenic filamentous microorganisms appear to be capable of producing similar mat morphologies as those seen in Oxyphotobacteria-dominated systems and commonly associated with

conical Precambrian stromatolites, and that the processes leading to the development of these features is more closely related with characteristics such as hydrology and cell morphology and motility.

Introduction

Stromatolites are "attached, lithified sedimentary growth structures, accretionary away from a point or limited surface of initiation" (Grotzinger and Knoll 1999). Behind this description lies a wealth of sedimentary structures with a record dating back over 3.7 billion years that may be one of the earliest indicators of life on Earth (Awramik 1992, Nutman et al. 2016). While some stromatolites form by abiotic processes (Grotzinger and Rothman 1996), many are constructed by microbial mats (e.g. Batchelor et al. 2004). Modern analog environments provide an opportunity to investigate key processes in stromatolite formation. These occur in environments such as hot springs (Berelson et al. 2011) and alkaline lakes (Petryshyn and Corsetti 2011) where microbial mats and sediments interact in the absence of eukaryotic grazing. The modern stromatolite analogs studied to date are dominated by Cyanobacteria (e.g. Pepe-Ranney et al. 2012, Reid et al. 2000), leading to the interpretation of ancient stromatolites as trace fossils of Cyanobacteria (e.g. Awramik 1992, Bosak et al. 2013). Cyanobacteria perform oxygenic photosynthesis, using light to fix carbon with water as an electron donor, producing oxygen as a byproduct. This process evolved from anoxygenic photosynthesis, which uses alternative electron donors and does not produce oxygen (Xiong et al. 2000). It is therefore reasonable to suspect that anoxygenic photosynthetic organisms supported the biosphere early in Earth history (Kharecha et al. 2005, Fischer and Knoll 2009, Ward et al. 2017b). Though it has

been proposed that anoxygenic microbes may have produced early stromatolites (Brock 1978, Bosask et al. 2007, Ward 2016), modern analogs have never been described without Cyanobacteria or filamentous algae contributing the mat fabrics. It remains to be demonstrated, therefore, that anoxygenic photosynthetic microbes—or even nonphotosynthetic microbes—can also produce stromatolites. If they can, it is possible that some or all of the oldest stromatolites record the activities of microbes other than Cyanobacteria. This would be consistent with the observation that oxygen did not accumulate in Earth's atmosphere until ~2.35 Gya (Bekker et al. 2004, Ward et al. 2016), more than a billion years after the earliest known biogenic stromatolites (Allwood et al. 2006).

An excellent candidate for testing whether anoxygenic microbes can contribute to the production of characteristic stromatolite morphologies are the Chloroflexi. Chloroflexi (i.e. Green Nonsulfur Bacteria) are a phylum of gliding, filamentous bacteria possessing a wide diversity of metabolisms and ecological roles, but are best known as photoheterotrophs (Overmann 2008). Chloroflexi have been shown to be diverse and abundant in a range of environments including soil (Costello and Schmidt 2006, Will et al. 2010), groundwater (Hug et al. 2013), marine sediment (Inagaki et al. 2003), wastewater (Miura et al. 2007, Kragelund et al. 2007), and even the human microbiome (Campbell et al. 2014). Despite their environmental richness revealed by culture-independent surveys, most described Chloroflexi belong to a few subclades isolated from hot springs (Yamada and Sekiguchi 2009). The best-described members of the Chloroflexi belong to the genus *Chloroflexus*, including *C. aggregans* (Hanada et al. 1995) and *C. aurantiacus* (Pierson and Castenholz

1974). *C. aggregans* and other members of the class Chloroflexia are metabolically versatile, capable of growth either as aerobic heterotrophs or anoxygenic phototrophs (Gupta et al. 2013). Phylogenetic analysis of the phototrophy genes of Chloroflexi suggests that anoxygenic photosynthesis in this group predates the evolution of oxygenic photosynthesis in Cyanobacteria (Xiong et al. 2000), implying that this group is ancient and therefore a good candidate for investigating questions of early Earth history. Perhaps most importantly, Chloroflexi possess a filamentous morphology and the ability for gliding motility. Although known modern stromatolite analogs are formed by Cyanobacteria, recent studies have suggested that oxygenic photosynthesis is not the driving factor in the formation of these structures, but instead motility and a filamentous morphology might be the most important factors (Shepard and Sumner 2010, Ward et al. 2014, Gong et al. 2014, Frantz et al. 2015). These preliminary studies suggest that Chloroflexi or other filamentous bacteria might be capable of forming stromatolites. If this is the case, the fact that modern stromatolites are formed by Cyanobacteria might represent ecological drivers that make Cyanobacteria more competitive in these environments than anoxygenic or nonphototrophic groups. Demonstrating the ability of Chloroflexi to develop stromatolite-like morphologies might help to resolve the paradox of stromatolites appearing in the rock record long before evidence for atmospheric oxygen by relaxing constraints on the timing of evolution of Cyanobacteria.

Here, we characterize cone-forming microbial mats from Nakabusa Onsen, a sulfidic hot spring located in Nagano Prefecture, Japan. At this locality, cone-forming microbial mats develop similar morphologies in both downstream Cyanobacteria-rich regions as well

as upstream regions in which high temperatures and sulfidic conditions select against oxygenic Cyanobacteria. We show here that upstream mat fabrics are developed by filamentous Chloroflexi, with no structural support by Cyanobacteria. These mats nonetheless retain similar morphologies to downstream mats composed of filamentous Cyanobacteria, demonstrating that conical morphologies can be developed by filamentous anoxygenic phototrophic bacteria. This suggests that the mechanisms responsible for development of cones are unrelated to oxygenic photosynthesis, and are instead controlled by factors like hydrology and fluid flow or by motility of filamentous microbes.

Materials and methods:

Geological context:

Nakabusa Onsen is located in the Japanese Alps near Azumino, Nagano, Prefecture, Japan. The hot spring outflow sampled here is located at 36.392429N, 137.748038E. Nakabusa Onsen is a sulfidic, moderately alkaline hot spring with source waters near 70°C (Kubo et al. 2011). Source waters of the spring are slightly alkaline (pH 8.5–9.0) and sulfidic (0.046–0.123 mM, Nakagawa et al. 2002, Nakagawa et al. 2003).

Sample collection:

Samples were collected and experiments were collected on 17 September 2015. Samples of two cone-forming microbial mats, CP1 and CP2, were collected for 16S and metagenomic sequencing as well as for microscopy. CP1 was located ~2 meters from the main outflow of the hot spring, was 48 °C. CP2 was located ~3 meters further downstream and was 32 °C.

Samples were collected using sterile forceps and spatulas (~0.25 cm^3 of material). Samples for sequencing were collected in triplicate and processed immediately in the field. Cells were lysed and DNA preserved in the field using Zymo Terralyzer BashingBead Matrix and Xpedition Lysis Buffer (Zymo Research, Irvine, CA). Cells were disrupted immediately by attaching tubes to the blade of a cordless reciprocating saw (Black & Decker, Towson, MD) and operating for 1 minute. Samples for microscopy were fixed immediately upon collection in 2% paraformaldehyde in pH 8 HEPES buffer.

Sequencing and analysis:

Following return to the lab, DNA was extracted and purified with a Zymo Soil/Fecal DNA extraction kit. The V4-V5 region of the 16S rRNA gene was amplified from each extract as well as negative controls using archaeal and bacterial primers 515F (GTGCCAGCMGCCGCGGTAA) and 926R (CCGYCAATTYMTTTRAGTTT) (Caporaso et al., 2012). DNA was quantified with a Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA) according to manufacturer's instructions following DNA extraction and PCR steps. All samples yielded PCR amplicons when viewed on a gel after initial pre-barcoding PCR (30 cycles). Duplicate PCR reactions were pooled and reconditioned for five cycles with barcoded primers. Samples for sequencing were submitted to Laragen (Culver City, CA) for analysis on an Illumnia MiSeq platform. Sequence data were processed using QIIME version 1.8.0 (Caporaso et al., 2010). Raw sequence pairs were joined and quality-trimmed using the default parameters in QIIME. Sequences were clustered into de novo operational taxonomic units (OTUs) with 99% similarity using UCLUST open reference clustering protocol (Edgar, 2010). Then, the most

abundant sequence was chosen as representative for each de novo OTU (Wang et al., 2007). Taxonomic identification for each representative sequence was assigned using the Silva-115 database (Quast et al., 2013) clustered at 97% similarity. Singletons and contaminants (OTUs appearing in the negative control datasets) were removed. 16S sequences were aligned using MAFFT (Katoh et al. 2002) and a phylogeny constructed using FastTree (Price et al. 2010). Summary statistics were calculated using scripts in QIIME and are reported at the 97% OTU similarity levels.

Imaging

Intact cones were fixed in 2% paraformaldehyde solution and embedded in resin for microtome sectioning and observation with fluorescent microscopy using photosynthetic pigment autofluorescence and morphology to distinguish taxa.

Results and Discussion:

16S amplicon sequencing, shotgun metagenomic sequencing, and optical and fluorescent microscopy all confirmed that Cone Pool 1 contained only trace Cyanobacteria, and was instead formed primarily by *Chloroflexus*. Cone Pool 2, in contrast, was made up primarily of filamentous Cyanobacteria with only trace Chloroflexi.

Microbial community composition

Relative abundance of microbial taxa in the Cone Pool mats were estimated via 16S amplicon sequencing. Results are summarized in Figure 2, showing the difference in relative abundance of taxa between the two mats.

Samples from CP1 recovered, on average, 4106 reads and 1073 OTUs (at the 97% cutoff), while CP2 recovered an average of 35547 reads and 6756 OTUs. Despite this

difference in sequencing coverage, the Goods Coverage metric of sampling depth averaged 0.78 for CP1 and 0.86 for CP2, reflecting fairly similar recovery of rare diversity. Dominant taxa were well represented in both datasets, though rare diversity was likely unrecovered from both sites.

Both communities were dominated by just a few phyla. The CP1 community was primarily made up of anoxygenic phototrophic and nonphototrophic Chloroflexi, as well as Thermotoga. The Chloroflexi phylum alone made up approximately one third of the reads form CP1, with Thermotogae and Proteobacteria (~20 and 10%, respectively) also abundant. No other phylum made up more than 5% of total reads. The CP2 community had a more even diversity. Chlorobi were the most abundant phylum (25%), with large numbers of Cyanobacteria (22%), Chloroflexi (16%), and Bacteroidetes (12%).

In CP1, the microbial community had low OTU-level diversity, with most taxa represented by only one or two dominant OTUs. The primary exception to this was the Thermotogae, which included four OTUs at 1% relative abundance or above, and *Chloroflexus*, with two OTUs. CP2, in contrast, had much higher strain-level diversity, with most taxa present as multiple OTUs at relatively even abundance (e.g. seven Chlorobi OTUs between 0.1 and 1% relative abundance or above).

Members of the Chlorobi were the most abundant taxon in CP2; this is despite the volumetric dominance of Cyanobacteria, potentially as a result of the large cell size of filamentous Cyanobacteria relative to Chlorobi or differential DNA extraction or amplification (e.g. Trembath-Reichert et al. 2016). The Chlorobi present at Nakabusa are basal members of the phylum, 91% similar to *Chloroherpeton thalassium* 35110. Given the

oxygen content of the mats in which they are found, these strains may be aerobic, a feature only described in basal members of the phylum which may record an ancestral trait (Liu et al. 2012, Stamps et al. 2014, Fischer et al. 2016).

In addition to the phototrophic *Chloroflexus* and *Roseiflexus,* both mats contained abundant nonphototrophic Chloroflexi, including Anaerolineae, Caldilineae, and *Herpetosiphon*. These taxa are typically described as filamentous heterotrophs (e.g. Yamada et al. 2006, Ward et al. 2015b). While the Anaerolineae have classically been described as obligate anaerobes (Yamada et al. 2006), genomic sequencing has revealed widespread genes for aerobic respiration in this class (e.g. Hemp et al. 2015a, Hemp et al. 2015b, Pace et al. 2015, Ward et al. 2015a), so the strains at Nakabusa may also be aerobic.

Preliminary optical and autofluorescence microscopy of embedded and sectioned intact cones from CP1 and CP2 has been performed, although rigorous analysis remains to be performed. This analysis confirmed that CP1 contained Chloroflexi as the primary filamentous component, with only unicellular Cyanobacteria, while CP2 was made up primarily of filamentous Cyanobacteria.

Filamentous microorganisms and cone formation

While overall diversity of the two mats is very distinct, focus should be drawn to the relative abundance of filamentous organisms in each mat. Filamentous organisms form cohesive fabrics in microbial mats, and are responsible for maintaining structural integrity, developing distinctive morphologies like cones, and trapping and binding sediment (Ward et al. 2014, Frantz et al. 2014). In CP1, the majority of filamentous organisms (as assessed by estimates of number and volume from microscopy as well as 16S sequencing reads)

were Chloroflexi, particularly the anoxygenic phototrophic genus *Chloroflexus*, along with lower abundances of *Roseiflexus,* Anaerolineae, and Caldilineae. Cyanobacteria were not abundant, and those present were primarily unicellular. CP2 contained much lower abundances of phototrophic Chloroflexi, and the filamentous community was instead composed predominantly of Anaerolineae and filamentous Cyanobacteria related to *Leptolyngbya* (a common member of hot spring microbial mats of similar temperatures, e.g. Roeselers et al. 2007, Bosak et al. 2012, Ward et al. 2017a). *Chloroflexus* and filamentous Cyanobacteria like *Leptolyngbya* are not closely related, having evolved filamentous morphologies and phototrophy independently. The nature of their photosynthetic pathways also differs in their reaction centers, carbon fixation pathways, and metabolic byproducts (Fischer et al. 2016, Shih et al. 2017). Despite these differences, the growth of these organisms has led to the development of similar cone morphologies in CP1 and CP2. This reflects the convergent evolution of gross mat morphology between different lineages of microbes, and that factors other than taxonomic affinity, such as hydrology and fluid flow or cell morphology or motility control overall mat morphology. By the same token, similar microbial communities can lead to different mat morphologies (e.g. Trembath-Reichert et al. 2016).

Conclusions

Demonstrating that Chloroflexi can form stromatolite-like structures has several key implications, but first and foremost it provides confirmation that stromatolites need not necessarily reflect the presence of Cyanobacteria, which will help resolve a current billion-year discordance between records of stromatolites and environmental oxygen (e.g. Bosak et

al. 2007, Ward et al. 2016).

While the cone-forming microbial mats described here are made up primarily of photoautotrophic microbes, it remains possible that stromatolites could be formed by nonphototrophic microbes, such as filamentous methanogens or sulfur oxidizers. Additional analog characterization or experiments with nonphototrophic filamentous organisms can help resolve this possibility.

More specifically, conical stromatolites are among the oldest in the rock record (e.g. Nutman et al. 2016), and no known process is known to be capable of producing conical stromatolites in the absence of microbial mats (Grotzinger and Rothman 1996, Batchelor et al. 2004). It's therefore crucial to understand the formation processes of these structures to better interpret their occurrences in deep time. Demonstrating that conical morphologies need not always involve Cyanobacteria (e.g. Bosak et al. 2009) opens the possibility that other processes control conical mat formation, such as phototaxis (Walter et al. 1976), diffusion limitation of nutrients (Petroff et al. 2010), or aggregation of motile filaments (Shepard and Sumner 2010, Ward et al 2014, Gong et al 2014).

It is worth noting that the microbial mats investigated here are not mineralized, and so have no chance of entering the rock record. However, this lack of mineralization is a result only of the geochemical details of the hot spring environment. Given the ubiquity of Chloroflexi in a range of environments (Yamada and Sekiguchi 2009), including hot springs where mineralization is occurring (Brock 1978, Ward et al. 2014), it is a reasonable assumption that appropriate environments for mineralized Chloroflexi stromatolites can be found.

Figure 1: Photographs of cone-forming microbial mats at Nakabusa Onsen. Above: CP1, a cone-forming microbial mat growing at 48°C, whose fabric was made up of filamentous Chloroflexi. Below: CP2, a cone-forming microbial mat growing at 32°C, whose fabric was made up of filamentous Cyanobacteria.

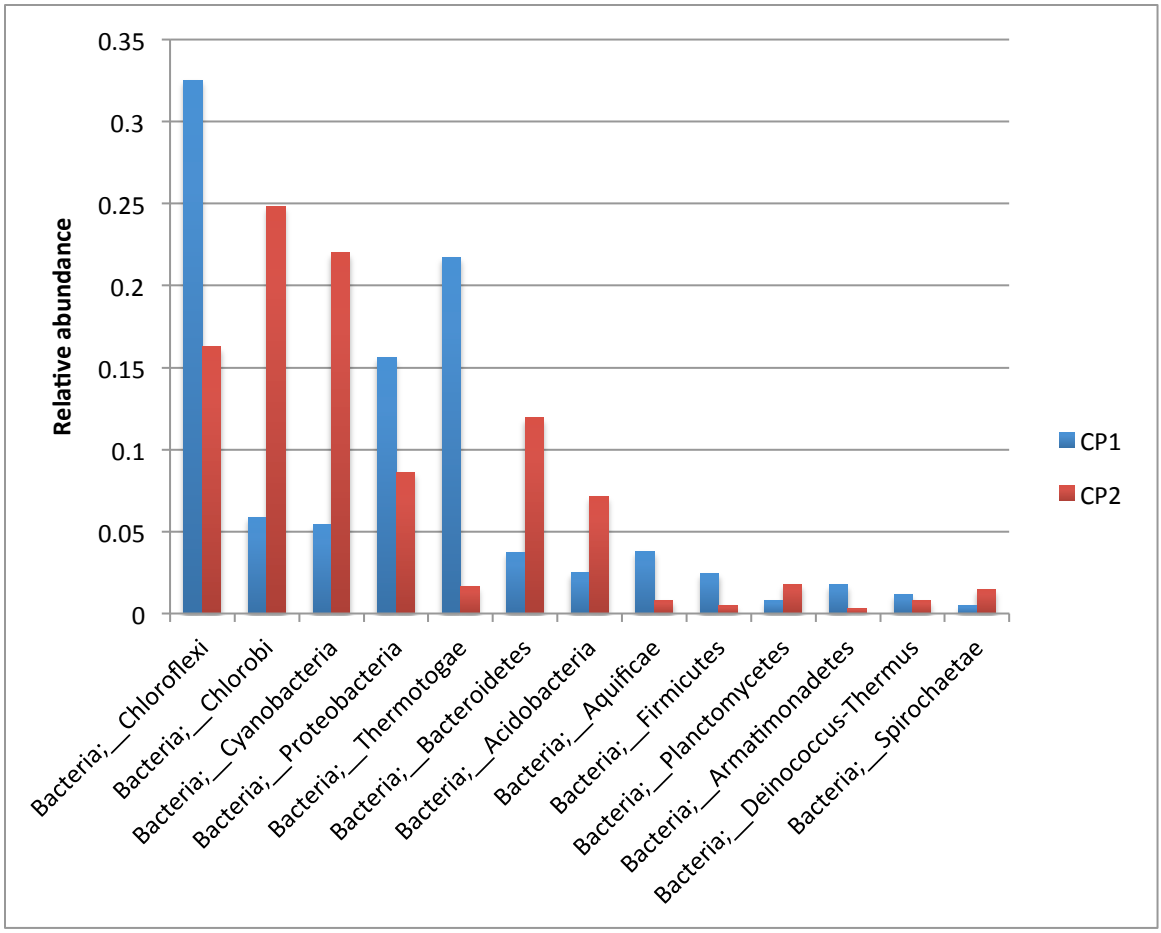


Figure 2: Relative abundance of most abundant taxa in Nakabusa samples clustered at the phylum level. Values presented are averaged across replicates within a site.

References:

1. Allwood, AC, MR Walter, BS Kamber, CP Marshall, and IW Burch. 2006. Stromatolite reef from the Early Archean era of Australia. Nature 441 pp714-718.
2. Awramik, SM. 1992. The oldest records of photosynthesis. Photosynthesis Research 33 pp75-89.
3. Batchelor, MT, RV Burne, BI Henry, and MJ Jackson. 2004. A case for biotic morphogenesis of coniform stromatolites. Physica A 337 pp319-326.
4. Bekker, A, et al. 2004. Dating the rise of atmospheric oxygen. Nature 427 pp117-120.
5. Berelson, WM, FA Corsetti, C Pepe-Ranney, DE Hammond, W Beaumont, and JR Spear. 2011. Hot spring siliceous stromatolites from Yellowstone National Park: assessing growth rate and laminae formation. Geobiology 9 pp411-424.
6. Bosak, T, et al. 2007. A Likely Role for Anoxygenic Photosynthetic Microbes in the Formation of Ancient Stromatolites. *Geobiology* 5(2) pp119–26.
7. Bosak, T, B Liang, MS Sim, and AP Petroff. 2009. Morphological record of oxygenic photosynthesis in conical stromatolites. PNAS 106(27) pp10939-10943.
8. Bosak, T, JWM Bush, MR Flynn, B Liang, S Ono, AP Petroff, and MS Sim. 2010. Formation and stability of oxygen-rich bubbles that shape photosynthetic mats. Geobiology 8 pp45-55.
9. Bosak, T. *et al*. Cyanobacterial diversity and activity in modern conical microbialites. *Geobiology* **10,** 384–401 (2012).

10. Bosak, T, AH Knoll, and AP Petroff. 2013. The meaning of stromatolites. Annu. Rev. Earth Sci. 41 pp21-44.
11. Brock TD (1978) Thermophilic microorganisms and life at high temperatures. Springer, New York
12. Campbell, AG, et al. 2014. Diversity and genomic insights into the uncultured *Chloroflexi* form the human microbiota. Environmental Microbiology 16(9) pp2635-2643.
13. Caporaso, JG, et al. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6 pp1621–1624.
14. Costello, EK and SK Schmidt. 2006. Microbial diversity in alpine tundra wet meadow soil: novel Chloroflexi from a cold, water-saturated environment. Environmental Microbiology 8(8) pp1471- 1486.
15. Fischer, WW and AH Knoll. 2009. An Iron Shuttle for Deepwater Silica in Late Archean and Early Paleoproterozoic Iron Formation. *Bulletin of the Geological Society of America* 121 pp222–35.
16. Fischer, W., Hemp, J. & Johnson, J. E. Evolution of Oxygenic Photosynthesis. *Annu. Rev. Earth Planet. Sci.* **44,** (2016).
17. Frantz, CM, V Petryshyn, and F Corsetti. 2015. Grain trapping by filamentous cyanobacterial and algal mats: implications for stromatolite microfabrics through time. *Geobiology* 13 pp409-423.
18. Gong, J, Z Zeng, and MM Tice. 2014. Evidence of microbial bioturbation and its unique records in clastic sediments. Geological Society of America Abstracts with

Programs. Vol. 46, No. 6, p.682.

19. Gupta, RS, P Chander, and S George. 2013. Phylogenetic framework and molecular signatures for the class Chloroflexi and its different clades; proposal for division of the class *Chloroflexi* class. Nov. into the suborder *Chloroflexineae* subord. nov., consisting of the emended family *Oschillochloridaceae* and the family *Chloroflexaceae* fam. nov., and the suborder *Roseiflexineae* subord. nov., containing the family *Roseiflexaceae* fam nov. Antonie van Leeuwenhoek 103 pp99-119.
20. Grotzinger, JP and DH Rothman. 1996. An abiotic model for stromatolite morphogenesis. Nature 383 pp423-425.
21. Grotzinger, JP and AH Knoll. 1999. Stromatolites in Precambrian carbonates: evolutionary mileposts or environmental dipstics? Annu. Rev. Earth Planet. Sci. 27 pp313-358.
22. Hanada, S, A Hiraishi, K Shimada, and K Matsuura. 1995. *Chloroflexus aggregans* sp. nov., a filamentous phototrophic bacterium which forms dense cell aggregates by active gliding movement. IJSEM 45(4) pp676-681.
23. Hemp J, **Ward LM**, Pace LA, Fischer WW. 2015a. Draft genome sequence of *Levilinea saccharolytica* KIBI-1, a member of the *Chloroflexi* class *Anaerolineae*. Genome Announc 3(6):e01357-15.
24. Hemp J, **Ward LM**, Pace LA, Fischer WW. 2015b. Draft genome sequence of *Ornatilinea apprima* P3M-1, an anaerobic member of the *Chloroflexi* class *Anaerolineae*. Genome Announc 3(6):e01353-15.

25. Hug, LA, CJ Castelle, KC Wrighton, BC Thomas, I Sharon, KR Frischkorn, KH Williams, SG Tringe, and JF Banfield. 2013. Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. Microbiome 1 pp1-22.
26. Inagaki, F, M Suzuki, K Takai, H Oida, T Sakamoto, K Aoki, KH Nealson, and K Horikoshi. 2003. Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. Appl. Environ. Microbiol. 69(12) pp7224-7235.
27. Kharecha, P, J Kasting, and J Siefert. 2005. A Coupled Atmosphere–ecosystem Model of the Early Archean Earth. *Geobiology* 3 pp53–76.
28. Kragelund, C, C Levantesi, A Borger, K Thelen, D Eikelboom, V Tandoi, Y Kong, J van der Waarde, J Krooneman, S Rossetti, TR Thomsen, ad PH Nielsen. 2007. Identitiy, abundance and ecophysiology of filamentous *Chloroflexi* species present in activated sludge treatment plants. FEMS Microbiol Ecol 59 pp671-682.
29. Kubo, K, K Knittel, R Amann, M Fukui, and K Matsuura. 2011. Sulfur-metabolizing bacterial populations in microbial mats of the Nakabusa hot spring, Japan. Systematic and Applied Microbiology 34(4) pp293-302.
30. Liu, Z. *et al*. Candidatus Thermochlorobacter aerophilum: an aerobic chlorophotoheterotrophic member of the phylum Chlorobi defined by metagenomics and metatranscriptomics. *ISME J*. **6,** 1869–1882 (2012).
31. Miura, Y, Y Watanabe, and S Okabe. 2007. Significance of *Chloroflexi* in performance of submerged membrane bioreactors (MBR) treating municipal

wastewater. Environ. Sci. Technol. 41 pp7787-7794.

32. Nakagawa, T., Fukui, M. (2002) Phylogenetic characterization of microbial mats and streamers from a Japanese alkaline hot spring with a thermal gradient. J. Gen. Appl. Microbiol. 48, 211–222.
33. Nakagawa, T., Fukui, M. (2003) Molecular characterization of community structures and sulfur metabolism within microbial streamers in Japanese hot springs. Appl. Environ. Microbiol. 69, 7044–7057
34. Nutman, A. P., Bennett, V. C., Friend, C. R. L., Van Kranendonk, M. J. & Chivas, A. R. Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. *Nature* 1–12 (2016). doi:10.1038/nature19355
35. Overmann, J. 2008. Green nonsulfur bacteria. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.
36. Pace LA, Hemp J, **Ward LM**, Fischer WW. 2015. Draft genome of *Thermanaerothrix daxensis* GNS-1, a thermophilic facultative anaerobe from the *Chloroflexi* class *Anaerolineae*. Genome Announc 3(6):e01354-15.
37. Pepe-Ranney, C, WM Berelson, FA Corsetti, M Treants, and JR Spear. 2012. Cyanobacterial construction of hot spring siliceous stromatolites in Yellowstone National Park. Environmental Microbiology 14(5) pp1182-1197.
38. Petroff, AP, MS Sim, A Maslov, M Krupenin, DH Rothman, and T Bosak. 2010. Biophysical basis for the geometry of conical stromatolites. PNAS 107(22) pp9956-9961.
39. Petryshyn, V and FA Corsetti. 2011. Analysis of growth directions of columnar

stromatolites from Walker Lake, western Nevada. Geobiology 9 pp425-435.

40. Pierson, BK and RW Castenholz. 1974. A phototrophic gliding filamentous bacterium of hot springs, Chloroflexus aurantiacus, gen. and sp. nov. *Archives of Microbiology* 100 pp5-24.
41. Reid, RP, PT Visscher, AW Decho, JF Stolz, BM Bebout, C Dupraz, IG Macintyre, HW Paerl, JL Pinckney, L Prufert-Bebout, TF Steppe, and DJ DesMarais. 2000. The role of microbes in accretion, lamination and early lithification of modern marine stromatolites. Nature 406 pp989- 992.
42. Roeselers, G. et al. Diversity of phototrophic bacteria in microbial mats from Arctic hot springs (Greenland). Environ. Microbiol. 9, 26–38 (2007).
43. Shepard, RN and DY Sumner. 2010. Undirected motility of filamentous cyanobacteria produces reticulate mats. Geobiology 8 pp179-190.
44. Shih, PM, LM Ward, J Hemp, and WW Fischer. 2017. Proterozoic origin of phototrophic Chloroflexi and the 3-hydroxypropionate bicycle. In prep.
45. Stamps, B. W., Corsetti, F. A., Spear, J. R. & Stevenson, B. S. Draft Genome of a Novel Chlorobi Member Assembled by Tetranucleotide Binning of a Hot Spring Metagenome. *Genome Announc*. **2,** 4669 (2014).
46. Trembath-Reichert E, **Ward LM**, Slotznick SP, Bachtel SL, Kerans C, Grotzinger JP, Fischer WW (2016) Gene sequencing microbial community analysis of mat morphologies, Caicos Platform, British West Indies, *Journal of Sedimentary Research* 86(6) pp629-636.

47. Walter, M. R., Bauld, J., and Brock, T. D., 1976, Microbiol- ogy and morphogenesis of columnar stromatolites (Conophyton, Vacerrilla) from hot springs in Yellow- stone Park, in Walter, M. R., ed., Stromatolites: Amster- dam, Elsevier, p. 273–310. Walter,

48. Ward, LM, Z Kerrigan, H Agic, M Juarez Rivera, VA Petryshyn, BW Stamps, HA Johnson, BS Stevenson, JR Spear, and FA Corsetti. 2014. Metagenomic insights into cone formation in hot spring microbial mats. Geological Society of America Abstracts with Programs. Vol. 46, No. 6, p.472.

49. **Ward LM**, Hemp J, Pace LA, Fischer WW. 2015a. Draft genome sequence of *Leptolinea tardivitalis* YMTK-2, a mesophilic anaerobe from the *Chloroflexi* class *Anaerolineae*. Genome Announc 3(6):e01356-15.

50. **Ward LM**, Hemp J, Pace LA, Fischer WW. 2015b. Draft genome sequence of *Herpetosiphon geysericola* GC-42, a nonphototrophic member of the *Chloroflexi* class *Chloroflexia*. Genome Announc 3(6):e01352-15

51. **Ward, L. M.,** Kirschvink, J. L. & Fischer, W. W. 2016. Timescales of Oxygenation Following the Evolution of Oxygenic Photosynthesis. Orig. Life Evol. Biosph. 46(1) pp51-65.

52. **Ward, LM,** A Idei, T Kakegawa, WW Fischer, and SE McGlynn. 2017a. Microbial diversity and iron oxidation at Okuoku-hachikurou Onsen, a Japanese hot spring analog of Precambrian iron formation. Geobiology, in review.

53. Ward, Lewis M, Birger Rasmussen, and Woodward W. Fischer. 2017b. Electron donor limitation of the biosphere before oxygenic photosynthesis. In preparation.

54. Will, C, A Thuermer, A Wollherr, H Nacke, N Herold, M Schrumpf, J Gutknecht, T Wubet, F Buscot, and R Daniel. 2010. Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. Appl. Environ. Microbiol. 76(20) pp6751-6759.

55. Xiong, J, WM Fischer, K Inoue, M Nakahara, and CE Bauer. 2000. Molecular evidence for the early evolution of photosynthesis. Science 289 pp1724-1730.

56. Yamada, T, H Imachi, A Ohashi, H Harada, S Hanada, Y Kamagata, and Y Sekiguchi. 2007. Bellilinea Caldifistulae Gen. Nov., Sp. Nov and Longilinea Arvoryzae Gen. Nov., Sp. Nov., Strictly Anaerobic, Filamentous Bacteria of the Phylum Chloroflexi Isolated from Methanogenic Propionate-Degrading Consortia."*International Journal of Systematic and Evolutionary Microbiology* 57 pp2299–2306.

57. Yamada, T and Y Sekiguchi. 2009. Cultivation of uncultured *Chloroflexi* subphyla: significance and ecophysiology of formerly uncultured *Chloroflexi* 'Subphylum I' with natural and biotechnological relevance. *Microbes Environ*. 24pp205–216.

A p p e n d i x 3

MICROBIAL DIVERSITY AND PRODUCTIVITY OF JINATA ONSEN, AN IRON-RICH INTERTIDAL HOT SPRING IN JAPAN

Lewis M. Ward, Airi Idai, Woodward W. Fischer, and Shawn E. McGlynn. In preparation.

Abstract

The redox history of Earth's atmosphere and oceans has varied dynamically through time, profoundly impacted by and in turn impacting life. The fluid Earth today is well oxygenated thanks to oxygenic photosynthesis by plants, algae, and Cyanobacteria, yet this state of affairs is unique to the Phanerozoic. Throughout the Proterozoic, the atmosphere contained at least some free oxygen while much of the oceans remained anoxic with varying amounts of free sulfide and ferrous iron. The interactions between these redox-active compounds and their role in biological productivity is not well constrained, partially as a result of the lack of appropriate process analogs for investigating the diversity and productivity of microorganisms supported by a range of combinations of free oxygen, iron, and sulfur compounds.

Hydrothermal environments, including terrestrial hot springs and near-shore hydrothermal vents, can contain diverse geochemical conditions that vary over short spatial scales thanks to the interactions between the oxygenated atmosphere, reducing hydrothermal fluids, and in some cases seawater.

One such process analog environment occurs at Jinata Onsen, on Shikinejima Island, Japan, where an intertidal, anoxic, iron- and hydrogen-rich hot spring mixes with

the oxygenated atmosphere and sulfate-rich seawater over short spatial scales, creating an enormous range of redox environments. Particular locations along the spring outflow support microbial communities and precipitated mineral textures potentially analogous to a range of early Earth environments, despite being separated by no more than a few meters.

Here, we characterize the geochemical conditions along the outflow of Jinata Onsen, as well as characterization of the microbial communities present via 16S amplicon metagenomic sequencing. Near the hot spring source, productivity is driven by oxidation of molecular hydrogen and ferrous iron by members of the Aquificales and Zetaproteobacteria, respectively, while downstream the microbial community transitions to being dominated by oxygenic Cyanobacteria. Cyanobacteria abundance and productivity dominates even at ferrous iron concentrations up to ~150 μM, challenging previous hypotheses of iron toxicity as a limitation to cyanobacterial expansion in the Precambrian ocean.

Background

Throughout Earth history, the metabolic opportunities available to life, and the resulting organisms and metabolisms responsible for driving primary productivity, have been shaped by the geochemical conditions of the atmosphere and oceans. Over the course of Earth's 4.56 billion year history, the redox state and overall geochemical conditions of the oceans have varied tremendously. The modern, sulfate-rich, well-oxygenated ocean that we see today is a relatively recent state, typical only of only the last few hundred million years (e.g. Lyons et al. 2014). For the first half of Earth history, until ~2.3 billion years ago (Gya), the atmosphere and oceans were anoxic (Johnson et al. 2014), and the oceans were

largely rich in dissolved iron but poor in sulfur (Walker and Brimblecombe 1985). Following the Great Oxygenation Event ~2.3 Gya, the atmosphere and surface ocean accumulated some oxygen, and the ocean transitioned into a stratified state with oxygenated surface waters and anoxic deeper waters, rich in either dissolved iron or sulfide (Poulton et al. 2010). This stratified ocean, though possibly variable, largely persisted for 1.5 billion years until the Neoproterozoic Oxygenation Event led to the ventilation of the deep oceans and the development of relatively modern ocean conditions (Sahoo et al. 2012). It is only under modern, well-oxygenated ocean conditions that Earth has supported the development of complex multicellular organisms like animals (Catling et al. 2005), and the biosphere was almost exclusively microbial throughout the earlier shifts in ocean chemistry.

As ocean conditions changed through the Archean and Proterozoic, the metabolic opportunities available to life also changed. The accumulation of oxygen enabled aerobic respiration and oxygen-fueled geochemical cycling of compounds like nitrogen and sulfur, while the abundance of compounds like ferrous iron and sulfide determined the potential productivity of anoxygenic phototrophs and lithotrophs. While we can gain some insight into the response of the biosphere to shifting ocean conditions over the course of Earth history by interpreting geochemical signatures in the rock record (e.g. the nitrogen isotope record of the evolution of the biogeochemical nitrogen cycle, Zerkle et al. 2017), an independent, complementary approach is to observe the activity of natural microbial communities in process analog environments which reflect some characteristics of ancient Earth environments. Many individual environments have been characterized that are

interpreted to be analogous to a particular period in Earth history; these include Lake Matano, in Indonesia, interpreted as being analogous to the ferruginous ocean (Crowe et al. 2008), Oku-Okuhachikurou Onsen in Akita Prefecture, Japan, similar to conditions just following the GOE (Ward et al. 2017), and the Black Sea, analogous to the stratified, anoxic- and sulfidic-at depth Proterozoic ocean (Scott et al. 2008). These analogs are each valuable in their own right, but the major differences between context at each site makes it difficult to isolate individual variables that lead to shifts in microbial community and productivity. An ideal test case for understanding the shifts of microbial productivity over the course of Earth history would be a system in which conditions of oxygen concentration and abundance of compounds such as sulfur and iron vary over short spatial scales under otherwise similar conditions. One such system occurs at Jinata Onsen, on Shikinejima Island, Tokyo Prefecture, Japan. At Jinata, an iron-rich hot spring emerges near a small bay, and mixes with seawater over the course of a few meters, quickly transitioning from an iron-rich and oxygen-poor condition, analogous to some conditions in the early Proterozoic, toward iron-poor and sulfate- and oxygen-rich conditions typical of the modern open ocean. Here, we characterize the geochemistry of this hot spring and describe the microbial community along the hot spring outflow as it mixes with ocean water via culture-independent 16S amplicon metagenomic sequencing. We show that the microbial community shifts from being dominated by iron- and hydrogen-oxidizing bacteria toward Cyanobacteria downstream. Furthermore, we show that Cyanobacteria become abundant even at high (>100 μM) Fe2+ concentrations, challenging interpretations of iron toxicity determining cyanobacterial ecology early in Earth history. Furthermore, we identify novel

high-temperature lineages of a number of microbial taxa, including iron-oxidizing Zetaproteobacteria.

Materials and Methods:

Geological context and sedimentology of Jinata:

Jinata Onsen is located at approximately 34.318N, 139.216E on the island of Shikinejima, Tokyo Prefecture, Japan. Shikinejima is part of the Izu Islands, a chain of volcanic islands which formed in the last few million years along the northern edge of the Izu-Bonin-Mariana Arc (Kaneoka et al. 1970). Shikinejima is formed of Late Paleopleistocene- to-Holocene non-alkaline felsic volcanics and Late-Miocene to Pleistocene non-alkaline pyroclastic volcanic flows (Figure 1).

The source water of Jinata Onsen emerges anoxic, iron-rich, and gently bubbling from the spring source (Figure 2). Temperatures at the source are ~62°C. Water emerges into the Source Pool, which has no visible microbial mat or biofilms. Surfaces are instead coated with a fluffy red precipitate, likely a poorly-ordered iron oxide phase such as ferrihydrite. The Source Pool has no mixing with seawater. Downstream, the spring water collects into a series of pools (Pool 1-3), which sequentially cool. Pool 1 is dominated by precipitated iron oxides, like the Source Pool, but also contains what appear to be microbial streamers that are mineralized and coated in iron oxides. Downstream pools (Pools 2 and 3) also mix with seawater during high tide. Samples were collected and temperatures were measured at high tide, reflecting the lowest temperatures experienced by microbes in the pools—at low

tide, hot spring input is dominant and temperatures rise. Surfaces in Pools 2 and 3 are covered in thick microbial mats. In Pool 2, the mat is coated in a layer of fluffy iron oxide similar to that in the source pool, with dense microbial mat below. Pool 3 contains only patchy iron oxides, with mostly exposed microbial mats displaying a finger-like morphology, potentially related to turbulent mixing during high tide. The Outflow is the outlet of a channel connecting Pool 2 to the bay. It is dominantly marine with a constant flow of spring water.

Sample collections:

Samples were collected from 5 sites at Jinata Onsen: the Source Pool, Pool 1, Pool 2, Pool 3, and the Outflow. Two samples were collected per site. At the Source Pool, Sample A was taken from the top of a ~10cm cobble and Sample B was collected from the bottom. In Pool 1, Sample A was taken of the yellow precipitate deeper in the pool, while Sample B was the redder, shallower section with abundant streamer-like structures. Pool 2 Sample A was an orange-ish mat collected from the side of a boulder. Pool 2 Sample B was a more cohesive greenish mat. Pool 3 Sample A was one of the finger-like mat structures. Sample B was a mottled, rough orange-green mat sample. At the Outflow, Sample A was taken from within the stream exiting Pool 2, while Sample B was taken of a mat within the swash zone.

Samples were collected as mineral scrapings of floc-y iron oxides upstream (Source Pool and Pool 1) and as samples of microbial mat downstream (Pools 2 and 3, and Outflow) using sterile forceps and spatulas (~0.25 cm^3 of material). Cells were lysed and DNA preserved in the field using a Zymo Terralyzer BashingBead Matrix and Xpedition

Lysis Buffer. Cells were disrupted immediately by attaching tubes to the blade of a cordless reciprocating saw and operating for 1 minute. Samples for geochemical analysis consisted of water collected via sterile syringe and filtered immediately through a 0.2 micron filter and gas bubbles collected from the source pool via funnel and sterile syringe and injected immediately into an anoxic serum vial to positive pressure.

Geochemical analysis:

Dissolved oxygen (DO), pH, and temperature measurements were performed *in situ* using an Exetech DO700 8-in-1 Portable Dissolved Oxygen Meter. Iron concentrations were measured using the ferrozine assay (Stookey 1970) following acidification with 40 mM sulfamic acid to inhibit iron oxidation by O_2 or oxidized nitrogen species (Klueglein and Kappler 2013). Ammonia/ammonium concentrations were measured using a TetraTest NH_3/NH_4^+ Kit with a Thermo Scientific Nanodrop 2000c spectrophotometer. Anion concentrations were measured via ion chromatography on a Shimadzu Ion Chromatograph equipped with a Shodex SI-90 4E anion column. Gas content of bubbles was determined via Shimadzu GC-14A gas chromatograph.

Microscopy.

Light microscopy images (Figure 4) were taken on a Zeiss Axio Imagers 2 Upright microscope (Zeiss, Germany) with 40□and 100x objective lens. Fluorescent images of DAPI (4=,6-diamidino-2-phenylindole)-stained samples were taken with an excitation wavelength below 395 nm and an emission wavelength between 420 and 470 nm. The autofluorescence of Cyanobacteria was detected by exposing the sample to a wavelength of between 395 and 440 nm and detecting the emission at a wavelength of 470 nm. Images of

DAPI fluorescence and autofluorescence were overlain using the FIJI software package (http://pacific.mpi-cbg.de).

Sequencing and analysis:

Following return to the lab, DNA was purified with a Zymo Soil/Fecal DNA extraction kit (Zymo Research, Irvine, CA). The V4-V5 region of the 16S rRNA gene was amplified from each extract as well as negative controls using archaeal and bacterial primers 515F (GTGCCAGCMGCCGCGGTAA) and 926R (CCGYCAATTYMTTTRAGTTT) (Caporaso et al., 2012). DNA was quantified with a Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA) according to manufacturer's instructions following DNA extraction and PCR steps. All samples yielded PCR amplicons when viewed on a gel after initial pre-barcoding PCR (30 cycles). Duplicate PCR reactions were pooled and reconditioned for five cycles with barcoded primers. Samples for sequencing were submitted to Laragen (Culver City, CA) for analysis on an Illumnia MiSeq platform. Sequence data were processed using QIIME version 1.8.0 (Caporaso et al., 2010). Raw sequence pairs were joined and quality-trimmed using the default parameters in QIIME. Sequences were clustered into de novo operational taxonomic units (OTUs) with 99% similarity using UCLUST open reference clustering protocol (Edgar, 2010). Then, the most abundant sequence was chosen as representative for each de novo OTU (Wang et al., 2007). Taxonomic identification for each representative sequence was assigned using the Silva-115 database (Quast et al., 2013) clustered at separately at 99% and at 97% similarity. Singletons and contaminants (OTUs appearing in the negative control datasets) were removed. 16S sequences were aligned using MAFFT (Katoh et al.

2002) and a phylogeny constructed using FastTree (Price et al. 2010). Alpha diversity was estimated using the Shannon Index (Shannon 1948) and Inverse Simpson metric (1/D) (Simpson 1949; Hill 1973) (Table 2). We used the UniFrac distance metric (Lozupone and Knight 2005) to assess the microbial community phylogenetic similarity of microbial communities (Table 3). All statistics were calculated using scripts in QIIME and are reported at the 99% and 97% OTU similarity levels. Multidimensional scaling (MDS) analyses and plots to evaluate the similarity between different samples and OHK environments were produced in R using the vegan and ggplot2 packages (R Core Team 2014, Oksanen et al. 2016, Wickham 2009) (Figure 5).

Metagenomes were sequenced from Pool 1 Sample A (JP1) and Pool 3 Sample A (JP3). DNA was quantified with a Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA) according to manufacturer's instructions following DNA extraction. Purified DNA was submitted to SeqMatic LLC (Fremont, CA) for library preparation and sequencing via Illumina HiSeq technology; a single lane of 2x100 sequencing was shared between these two samples and two from another project. Raw sequences were assembled with MegaHit v. 1.02 (Li et al. 2016). Coverage was extracted using bbmap (Bushnell 2016) and samtools (Li et al. 2009). Genes of interest (e.g. coding for ribosomal, photosynthesis, iron oxidation, and electron transport proteins) were identified from assembled metagenomic data locally with BLAST+ (Camacho et al. 2008), aligned with MUSCLE (Edgar 2004), and alignments manually curated in Jalview (Waterhouse et al. 2009). Phylogenetic trees were calculated using RAxML (Stamakis 2014) on the Cipres science gateway (Miller et al. 2010).

Results

Geochemistry

The gas bubbling at the source was determined to contain CO_2, CH_4, and H_2. Source waters were slightly enriched in chloride relative to seawater (~23.2 g/L) but depleted in sulfate (~1.63 g/L) but approached seawater concentrations downstream as mixing increased.

Geochemistry measurements of Jinata source water are summarized in Table 1, while geochemical gradients along the stream outflow are summarized in Table 2. Water emerging from the source was 62°C, very low in dissolved oxygen (~0.15 mg/l), had a pH 5.4, and contained substantial concentrations of dissolved iron (~250 μM Fe^{2+}). After emerging from the source pool, the spring water exchanges gases with the air due to mixing associated with water flow and gas ebullition, and DO rose to 1.24 mg/L at the surface of the source pool. As water flows downstream from the source pool, it cools slightly, exchanges gases with the atmosphere, and intermittently mixes with seawater below Pool 1.

Sequencing

In total, we recovered 456,737sequences from the 10 samples at Jinata (Table 3). Reads per sample ranged from 26,057 Source Pool Sample A to 97,445 for Pool 1 Sample A (median 43,331, mean 45,673, and standard deviation 19,568). Assessment of sampling depth was estimated using Good's Coverage (Good 1953). On average, 74% of the microbial community was recovered from Jinata samples at the 99% OTU level based on

the Good's Coverage statistic (ranging from 54% coverage in the Outflow Sample A to 85% in the Pool 1 Sample A) and 87% at the 97% OTU level (74% for the Outflow Sample A to 94.5% for the Pool 1 Sample B).

Samples from the same site are highly similar, and adjacent sites (e.g. Source and Pool 1, Outflow and Pool 3) show significant similarity. However, there is essentially no overlap in diversity between distant samples (e.g. Source and Outflow).

The assembled contigs from JP1 totaled 334 MB in 331011 contigs. The Pool 3 assembly was 136 MB in 126181 contigs. In order to normalize relative abundance of genes in these two datasets despite their very different sizes, abundance of functional genes were normalized to the number of rpoB genes recovered in each metagenome (sum of distinct sequences assembled multiplied by their coverage). Like the 16S gene, RpoB is a highly conserved, vertically-inherited gene useful for taxonomic identification of organisms, but has the added advantage that it is only known to occur as a single copy per genome (Case et al. 2007). Approximately 11.5x more RpoB genes were recovered in the Pool 1 metagenome than that from Pool 3.

Discussion

Relative abundances of microbial taxa as revealed by 16S surveys can be useful for predicting metabolisms driving geochemical cycles at Jinata (Table 4). In particular, the contributions of various hydrogen- and iron-oxidizers and phototrophs to primary productivity along the spring path can be estimated due to these metabolisms being fairly well conserved within bacterial taxa (e.g. Emerson et al. 2010, Chan et al. 2016). Analysis

of the most abundant taxa at Jinata revealed significant roles for organisms driving aerobic iron and hydrogen oxidation upstream, while Cyanobacteria dominate downstream.

Despite sequencing to relatively high depth (>18000 reads per sample), rarefaction analysis shows that there is still substantial unsequenced diversity at Jinata (Supplemental Figure 1). This likely reflects rather uneven diversity, as >50% of reads at most sites are made up of the 10 most abundant taxa (Table 4). Additional sequencing is therefore likely to reveal additional diversity of rare taxa.

Because only two samples from Jinata were sequenced for shotgun metagenomics, these data cannot be used to interpret the fine scale gradient shown by 16S amplicon sequencing; however, metagenomic results can be used to interpret broad trends between the iron- and hydrogen-dominated upstream sections of the hot spring and the more oxic, lower-iron sections downstream. In particular, the relative abundance of functional metabolic genes associated with processes such as iron oxidation can be compared between these two sites to reinforce interpretations from the 16S data of the metabolic potential of communities in these sites.

Iron and hydrogen oxidation

The hot spring water emerging at the Source Pool at Jinata contains abundant dissolved Fe^{2+} and is continuously bubbling with H_2, and these highly favorable electron donors appear to fuel productivity and determine the microbial community upstream, particularly in the Source Pool and Pool 1 where microbial mats are not well developed.

The most abundant organisms in the Source Pool are members of the Aquificae family Hydrogenothermaceae. This family of marine thermophilic lithotrophs is capable of

both iron and hydrogen oxidation (Takai and Nakagawa 2014), and may be utilizing either or both of these electron donors at Jinata. The seventh most abundant OTU in the Source Pool samples is a novel sequence 89% similar to a strain of *Persephonella*, with relatives found in an alkaline hot spring in Papua New Guinea. *Persephonella* is a genus of thermophilic, microaerophilic hydrogen oxidizing bacteria within the Hydrogenothermaceae (Götz et al. 2002).

The other most abundant organisms near the source are Zetaproteobacteria, related to the neutrophilic, aerobic iron-oxidizing *Mariprofundus* common in marine systems (Emerson et al. 2007). Zetaproteobacteria and Hydrogenothermaceae together made up ~30-65% of sequences in the Source Pool and Pool 1, and so appear to be both numerically dominant as well driving the base of productivity in these environments.

The abundance of Hydrogenothermaceae drops off significantly to less than 1% of sequences downstream of Pool 1 once microbial mats become well developed, but Zetaproteobacteria continue to make up a few percent of reads until the Out Flow. This suggests that shifts in Zetaproteobacteria relative abundance may be due more to the increase in abundance of other organisms, rather than a drop in the number of Zetaproteobacteria or their ability to make a living oxidizing iron. This is consistent with the significant abundance and continual decline in iron concentrations along the flow path of the hot spring.

Members of the Mariprofundaceae have been observed to have an upper growth temperature of 30°C (Emerson et al. 2010). Jinata, with abundant Zetaproteobacteria found at temperatures up to 63 degrees, may represent a unique high-temperature environment for

these organisms. In particular, the third most abundant OTU in the Source Pool and Pool 1 sample A is an unknown sequence which is 92% identical to an uncultured Zetaproteobacteria sequence from a shallow hydrothermal vent in Papua New Guinea (Meyer-Dombard et al. 2013). This sequence likely represents a novel strain of high-temperature iron-oxidizing Zetaproteobacteria.

Members of several phototrophic Proteobacteria lineages were found at Jinata, though not at high relative abundances (<2% of sequences each). These include the Gammaproteobacteria family Chromatiales as well as the Alphaproteobacteria families Rhodospirillales and Rhodobacteraceae. Members of these clades are capable of photoautotrophy using diverse electron donors, including sulfur compounds, H_2 and Fe^{2+} (Imhoff 2014, Pujalte et al. 2014, Baldani et al. 2014), and so it is unclear which if any electron donor is fueling productivity by these organisms at Jinata.

The diversity of iron oxidizing bacteria at Jinata is very different than in other iron-rich hot springs. For example, at Oku-Okuhachikurou Onsen in Akita Prefecture, Japan, iron oxidation is driven primarily by the Gammaproteobacteria *Gallionella (*Ward et al. 2017), while at Chocolate Pots hot spring at Yellowstone National Park iron oxidation is driven primarily by oxygen produced *in situ* by Cyanobacteria (Pierson et al 1999). This may be related to selection by the highly saline water, or biogeographically by access to the ocean, as Zetaproteobacteria are typically found in marine settings, particularly in deep ocean basins associated with hydrothermal iron sources (Emerson et al. 2010). Similarly to Oku-Okuhachikurour Onsen, Jinata supports only limited biomass in regions dominated by iron oxidizers (Ward et al. 2017), reflecting the shared biochemistry and bioenergetic

challenges of neutrophilic iron oxidation by *Gallionella* and Zetaproteobacteria (Kato et al. 2015, Bird et al. 2011).

The relative abundance of genes associated with iron oxidation in the Pool 1 and Pool 3 metagenomes reflect the organismal abundance revealed by 16S amplicon sequencing. Diversity and coverage of foxE, mobB, and pioA—associated with iron oxidation in *Rhodobacter*, *Mariprofundus*, and *Rhodopseudomonasi*, respectively—were present at much higher abundance in the Pool 1 metagenome than Pool 3 (mobB and pioA at 4-5x higher relative abundance in Pool 1, while foxE was of similar abundance in Pool 1 to pioA but absent from the Pool 3 metagenome).

Cyanobacteria

Cyanobacteria are nearly absent from near the source pool, but begin to appear around Pool 1 and become dominant starting in Pool 2. The most abundant Cyanobacteria present are predominantly members of Subsection III, Family I. This group includes *Leptolyngbya*, a genus of filamentous non–heterocystous Cyanobacteria that has appeared in other hot springs of similar temperatures (e.g. Ward et al. 2017, Roeselers et al. 2007, Bosak et al. 2012). Cyanobacteria, performing oxygenic photosynthesis, appear to dominate primary productivity in downstream regions of the hot spring, and the filamentous morphology of the strains present here allow them to contribute to the cohesive fabric of the microbial mat.

In the Out Flow samples, Chloroplast sequences become abundant, most closely related to the diatom *Melosira*. Algae are at very low abundance upstream of the Out Flow,

potentially inhibited by high temperatures, high iron concentrations, or other characteristics of the hot spring water, but the higher seawater influence at the Out Flow creates a more permissive environment.

Cyanobacteria are sometimes underrepresented in iTag datasets as a result of poor DNA yield or amplification biases (e.g. Parada et al. 2015, Trembath-Reichert et al. 2016), but the low abundance of Cyanobacteria near the Source Pool was confirmed by fluorescent microscopy, in which cells displaying cyanobacterial autofluorescence were observed abundantly in samples from the downstream samples but not in the Source Pool (Figure 4).

Methane cycling

Methane is present in the bubbles emerging from the Source Pool at Jinata. No sequences associated with known methanogens were recovered from Jinata, suggesting that the methane found here may have an abiotic thermal source. This presence of methane does however provide a niche for methane oxidizing microbes, particularly aerobic methanotrophic bacteria. An abundant 16S sequence in the Source Pool and Pool 1 is most similar (94%) to that from *Methylomarinovum caldicuralii*, a thermophilic aerobic methanotroph from a hydrothermal vent in Japan (Hirayama et al. 2014), and so may represent a related strain involved in methane oxidation at Jinata. Related species in the *Methylothermaceae* family have recently been shown to also be capable of denitrification,

potentially coupled to methane oxidation (Skennerton et al. 2015) raising the possibility that members of this clade are more metabolically versatile than expected.

Anaerobic respiration

The amount of sulfur cycling occurring at Jinata likely varies significantly along stream. This is reflected in the relative abundance of common sulfate reducing bacteria such as the Deltaproteobacterium *Desulforomonas*, which increases greatly downstream. Sulfate is abundant throughout Jinata, but more thermodynamically favorable electron donors like oxygen and iron oxide likely make sulfate reduction unfavorable except within dense microbial mats. Furthermore, the presence of high concentrations of dissolved iron upstream make sulfide unstable, as it will rapidly precipitate as iron sulfide minerals like pyrite, reducing the ability of the microbial community to cycle sulfur effectively—conditions analogous to those in the dominantly ferruginous Archean ocean (Walker and Brimblecombe 1985).

The Deferribacteres phylum is present at up to ~2% relative abundance in some samples. The Deferribacteres is made up of thermophiles with diverse anaerobic respiratory pathways, including the reduction of iron, nitrate, or manganese (Garrity et al. 2001). Given the abundance of poorly ordered iron oxides like ferrihydrite at Jinata, iron reduction is a favorable anaerobic respiration pathway, creating a likely niche for these organisms.

Nitrogen cycling

Ammonia is fairly abundant in the source water of Jinata (~1.56 mg/L), and is still detectable by the outflow (~0.56 mg/L), suggesting that nitrogen is not a limiting nutrient in this hot spring. Some biological nitrogen cycling may be occurring, as evidenced by the

presence of sequences associated with nitrifying microbes. This includes Marine Group 1 Thaumarchaeota (up to 3% abundance in Pool 2), a clade made up of ammonia oxidizing archaea such as *Nitrosopumilus maritimus* (Pester et al. 2011). The family Nitrospiraceae in the Nitrospira phylum was present at up to ~2% abundance in upstream samples at Jinata. The Nitrospiraceae include the nitrite oxidizing *Nitrospira*, but the family is metabolically diverse and also includes aerobic iron oxidizers and both heterotrophic and hydrogenotrophic sulfate reducers (Daims 2014). However, the three most abundant OTUs affiliated with the Nitrospiraceae at Jinata are most similar to *Thermodesulfovibrio*, thermophilic, hydrogenotrophic or heterotrophic sulfate reducing bacteria (Daims 2014). A smaller number of 16S reads were associated with the *Nitrospira*, primarily recovered from Pool 1 Sample A; this is consistent with substantial coverage of nitrite oxidoreductase genes recovered from the metagenome collected from this site. The spike in relative abundance of *Nitrospira* where Pool 1 flows into Pool 2 may be a result of cooling of the hot spring water across the upper temperature limit of these organisms, as the upper temperature limit for cultured members of this genus is 52°C (Lebedeva et al. 2011).

Other organisms

Members of Subgroup 22 of the Acidobacteria phylum were abundant at Jinata. The Acidobacteria phylum is subdivided into 26 subgroups (Barns et al. 2007), yet only six contain well-characterized members (Kielak et al. 2016). This does not include Subgroup 22, and so it is unknown what role Acidobacteria may play at Jinata, but the abundance of these organisms at Jinata may make this hot spring a valuable resource for their characterization.

Members of the Chloroflexi class Anaerolineae are common throughout Jinata with the exception of the Outflow. The Anaerolineae have generally been isolated as obligately anaerobic heterotrophs (e.g. Sekiguchi et al. 2003, Yamada et al. 2006), but genome sequencing has revealed the capacity for aerobic respiration in diverse members of this clade (e.g. Hemp et al. 2015ab, Pace et al. 2015, Ward et al. 2015a). Furthermore, phototrophy has been described in a close relative of this class from a Yellowstone National Park metagenome (Klatt et al. 2011), further hindering attempts to interpret the metabolism and ecology of members of this genome from 16S data alone.

Additionally, members of the Chloroflexi class Caldilineae were present at up to ~1% abundance at Jinata. Members of the Caldilineae have previously been isolated form intertidal hot springs in Iceland (Kale et al. 2013) and Japanese hot springs (Sekiguchi et al. 2003) and placed into a separate class in the Chloroflexi sister to the Anaerolineae (Yamada et al. 2006). Characterized organisms in this class are filamentous, anaerobic, or facultatively aerobic heterotrophs and therefore may play a role in degrading biomass within low-oxygen regions of microbial mats.

Members of the Bacteroidetes phylum are abundant at Jinata, particularly in Pool 2 and Pool 3 where cyanobacterial mats are well developed. The Bacteroidetes sequences found are primarily members of the Sphingobacterales families Saprospiraceae and Chitinophagaceae. Saprospiraceae are commonly involved in the breakdown of complex organic matter in marine environments (McIllroy and Nielsen 2014), while members of the Chitinophagaceae are known to degrade biopolymers like cellulose and chitin (Rosenberg 2014), suggesting a possible role for these organisms in the degradation of extracellular

polymeric substances in the microbial mats at Jinata. The Verrucomicrobia family Opitutaceae is also abundant in microbial mats at Jinata, and is typically associated with the degradation of recalcitrant heteroppolysaccharides in a variety of contexts including marine and hot spring environments (Rodrigues and Isanapong 2014) and may fill a similar niche.

The Planctomycete family Phycisphaeraceae was also at relatively high abundance in Pool 2 and Pool 3. This lineage was first isolated from marine algae on Mikura-jima, part of the Izu Islands of which Shikinejima is a part (Fukunaga et al. 2009). Additional strains of Phycisphaeraceae have since been isolated in association with marine algae (Yoon et al. 2014), suggesting that Phycisphaeraceae at Jinata may also occur in association with algae at Jinata.

Members of the WS3 phylum, also known as Latescibacteria, are closely related to the Bacteroidetes as part of the CFB superphylum, and have been described as anaerobic fermenters that break down algal polysaccharides (Youssef et al. 2015). Members of the WS3 phylum may play a similar role at Jinata, but are most abundant in Pool 1, where microbial mats are absent and algae are at very low abundance, so it is possible that they fill a different niche.

Conclusions

Jinata Onsen is a novel environment supporting strong geochemical gradients over short spatial scales. The transition from low-oxygen, iron- and hydrogen-rich hot spring source water to oxygen-rich ocean water takes place over just a few meters, and results in an almost complete change in microbial community. This system is significant for its

relevance as a process analog for environments through Earth history and potentially habitable environments in Mars' past, and for its utility as environment to investigate iron-tolerant Cyanobacteria.

Over the course of ~1.5 billion years in the Proterozoic eon, the ocean underwent major geochemical and redox shifts from anoxic, iron-rich Archean oceans to its modern oxygen- and sulfate-rich state (Lyons et al. 2014). This major change in ocean geochemistry was reciprocally impacted by changes in the organisms and metabolisms used to drive primary production and other geochemical cycles. Through the Archean eon, the ocean was anoxic and iron-rich, and productivity is thought to have been driven primarily by iron- and hydrogen-oxidizing microbes (Kharecha et al. 2005, Canfield et al. 2006). At this time, and during the beginning of the Proterozoic, oxygenic Cyanobacteria were absent or unproductive (Fischer et al. 2016). Following the Great Oxygenation event ~2.3 billion years ago, oxygen accumulated to low but non-negligible levels in the atmosphere, and the surface ocean became oxygenated though the deep ocean remained anoxic and was variably rich in either free sulfide or ferrous iron (Poulton et al. 2010). Cyanobacteria and aerobic metabolisms became more significant, but anaerobic and anoxygenic organisms continued to contribute to productivity (Brocks et al. 2005, Johnston et al. 2009). Finally, during the Neoproterozoic, atmospheric oxygen reached modern levels and the deep ocean became well oxygenated (Saito et al. 2012). Since this time, marine productivity has been dominated by Cyanobacteria and eukaryotic algae (Geider et al. 2001).

These shifts in open ocean conditions took place over ~1.5 billion years, yet a full range of analogous conditions are recapitulated in some ways over a spatial scale of just a few meters at Jinata. The Source Pool, with its low oxygen, abundant ferrous iron and molecular hydrogen, and productivity dominated by lithotrophs and anoxygenic phototrophs in the absence of Cyanobacteria, is comparable to conditions in the earliest Paleoproterozoic, leading up to the Great Oxygenation Event. The gradual increase of oxygen and sulfate concentrations and loss of iron and hydrogen going downstream is comparable to middle Proterozoic time, when Cyanobacteria and anoxygenic phototrophs coexisted in a redox-stratified ocean. Finally, with a transition to fully oxygenated conditions and modern marine chemistry and the appearance of productive eukaryotic algae, the Out Flow at Jinata is representative of relatively modern ocean conditions that first developed in the Neoproterozoic. While other analog systems exist for each of these time periods (e.g. ferruginous Lake Matano in Indonesia as an Archean analog, Crowe et al. 2008, and Oku-Okuhachikurou Onsen in Japan as a Paleoproterozoic analog, Ward et al. 2017), these systems are incredibly disparate and challenging to compare directly. As a result of Jinata reflecting aspects of such a long span of Earth history over such a short spatial scale, this system allows comparison of very different geochemical contexts while avoiding extraneous confounding variables that plague comparisons of more disparate analog field sites, such as differences in biogeography, nutrient availability, or extreme differences in salinity or pH.

By observing trends in microbial diversity and community function across this gradient in geochemical conditions at Jinata, we can make predictions about how the

biosphere on the early Earth may have functioned. By observing, for instance, the relative productivity and abundance of aerobic iron- and hydrogen-oxidizers relative to anoxygenic phototrophs in the Source Pool we may be able to make predictions about the relative contribution of these metabolisms to the deposition of Banded Iron Formations through Earth history (e.g. Konhauser et al. 2002), while comparison of the net community productivity between the Source Pool and downstream regions we can predict how global productivity may have changed over Proterozoic time as oxygenic phototrophs began to dominate. The proximal cause of the GOE, and the transition toward a more oxygen-rich and iron-poor ocean, is contested; this event may record the late evolution of oxygenic Cyanobacteria (e.g. Ward et al. 2016, Shih et al. 2016, Fischer et al. 2016), or may reflect a delayed expansion of cyanobacterial productivity due to geochemical or geological conditions (e.g. Kasting 2013). Similarly, the Neoproterozoic saw a shift from productivity driven by Cyanobacteria and anoxygenic phototrophs to Cyanobacteria and eukaryotic algae (Johnston et al. 2009). Though again, whether the shift in dominant primary producers at this time was a result of the late evolution of eukaryotic algae (e.g. Shih and Matzke 2013) or due to a shift in geochemical conditions (e.g. Anbar and Knoll 2002) is unclear. Characterization of the factors controlling dominant primary producers along the flow path of Jinata could help build hypotheses for these historical transitions, as Jinata provides a case where the full suite of primary producers area present but a clear succession from hydrogen- and iron-driven metabolisms, to Cyanobacteria, to eukaryotic algae occurs along the flow path of the hot spring. Constraining what factors drive this transition—oxygen or iron concentrations, trace element abundance, or other factors that may or may

not be relevant to the Proterozoic ocean—may help to predict the drivers for historical shifts in productivity as well.

Jinata is also a valuable resource for investigating the ecology and mineralogical records of iron-oxidizing Zetaproteobacteria in an environment which is much more accessible than other Zetaproteobacteria-rich systems such as those at Loihi Seamount (e.g. Chan et al. 2016). Characterization of Zetaproteobacteria *in situ* is particularly valuable for characterizing the production and preservability of mineralogical signatures of these organisms. The iron oxides precipitated as a byproduct of their metabolism can retain distinctive morphologies, including "stalks" produced by Zetaproteobacteria like *Mariprofundus* (Chan et al. 2011). These iron microfossils can be preserved over geological timescales and so are useful for interpreting the role of iron oxidizing bacteria in past environments (Krepski et al. 2013), and may help diagnose what if any role aerobic iron oxidizing bacteria like the Zetaproteobacteria played in the deposition of Proterozoic Banded Iron Formations (Chan et al. 2016b).

Furthermore, our observations at Jinata are relevant for interpretations of the environmental constraints on cyanobacterial productivity in the Archean and early Proterozoic when dissolved ferrous iron is thought to have been abundant in the oceans. Previously, it has been suggested that high ferrous iron concentrations are toxic to Cyanobacteria, greatly reducing their potential for productivity under ferruginous conditions that may have persisted through much of the Archean era (Swanner et al. 2015). The high rates of cyanobacterial productivity observed at Jinata under high iron concentrations suggest that Cyanobacteria can adapt to ferruginous conditions, and

therefore iron toxicity is unlikely to inhibit Cyanobacteria over geological timescales. In fact, the iron concentrations observed at Jinata are much higher (150-250 μM) than predicted for the Archean ocean (<120 μM, Holland 1984) or observed at other iron-rich hot springs (~100-200 μM, Pierson et al. 1999, Ward et al. 2017), making Jinata an excellent test case for determining the ability of Cyanobacteria to adapt to high iron concentrations. Culture-based experiments may be useful to determine whether Jinata Cyanobacteria utilize similar strategies to other iron-tolerant strains (e.g. the *Leptolyngbya*-relative *Marsacia ferruginose*, Brown et al. 2010) or whether Jinata strains possess unique adaptations that allow them to grow at higher iron concentrations than known for other environmental Cyanobacteria strains, as well as whether the paucity of Cyanobacteria in the Source Pool and Pool 1 are due to the >260 μM iron concentrations or a result of other factors like high temperatures.

Finally, the dynamic abundances of redox-active compounds oxygen, iron, hydrogen, and sulfate at Jinata may not only be analogous to conditions on the early Earth, but may have substantial relevance for potentially habitable environments on Mars as well. Early Mars is thought to have supported environments with metabolic opportunities provided by the redox gradient between the oxidizing atmosphere and abundant electron donors such as ferrous iron and molecular hydrogen sourced from water/rock interactions (e.g. Hurowitz et al. 2010). Uncovering the range of microbial metabolisms present under the environmental conditions at Jinata, and their relative contributions to primary productivity, may therefore find application to predicting environments on Mars most able to support productive microbial communities.

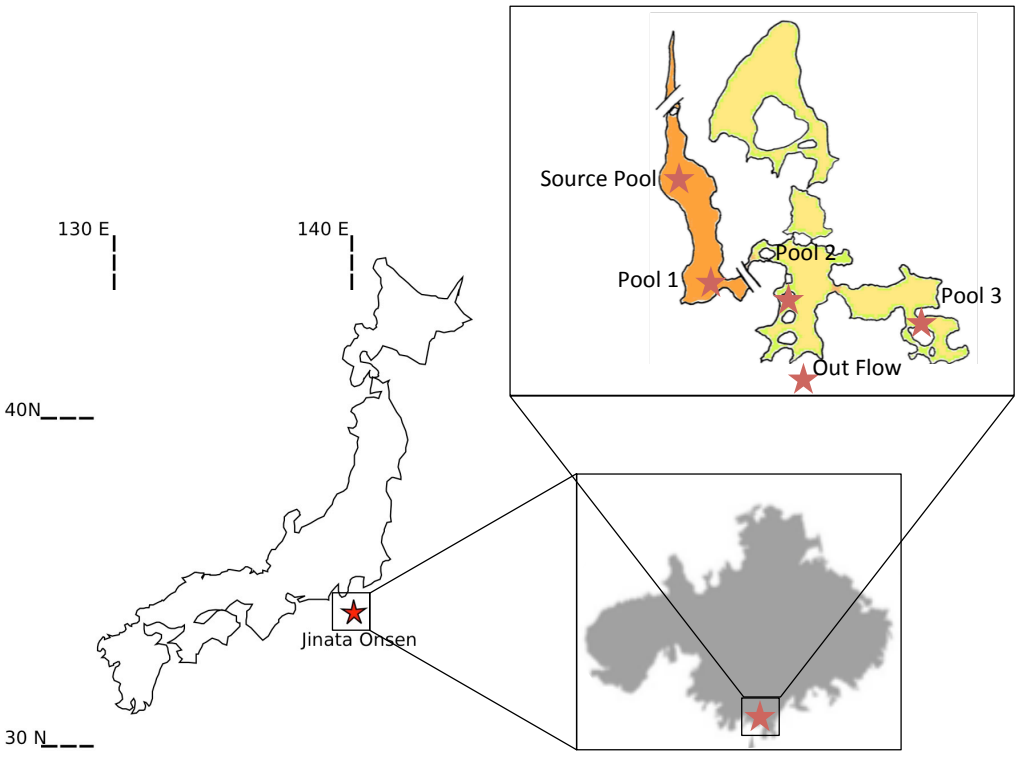


Figure 1:

Location of Jinata Onsen on Shikinejima Island, Japan, and inset overview sketch of field site with sampling localities marked.

Figure 2:

Representative photos of Jinata. A) Panorama of field site, with source pool on left (Pool 1 below), Pool 2 and 3 in center, and Out Flow to bay on right. B) Undistorted view north up the canyon. C) Undistorted view south toward bay, overlooking Pool 2. D) Source pool, coated in floc-y iron oxides and bubbling with gas mixture containing H2, CO2, and CH4. E) Pool 2, with mixture of red iron oxides and green from Cyanobacteria-rich microbial mats. F) Close up of textured microbial mats in Pool 3. G) Close up of Out Flow, where hot spring water mixes with ocean water.

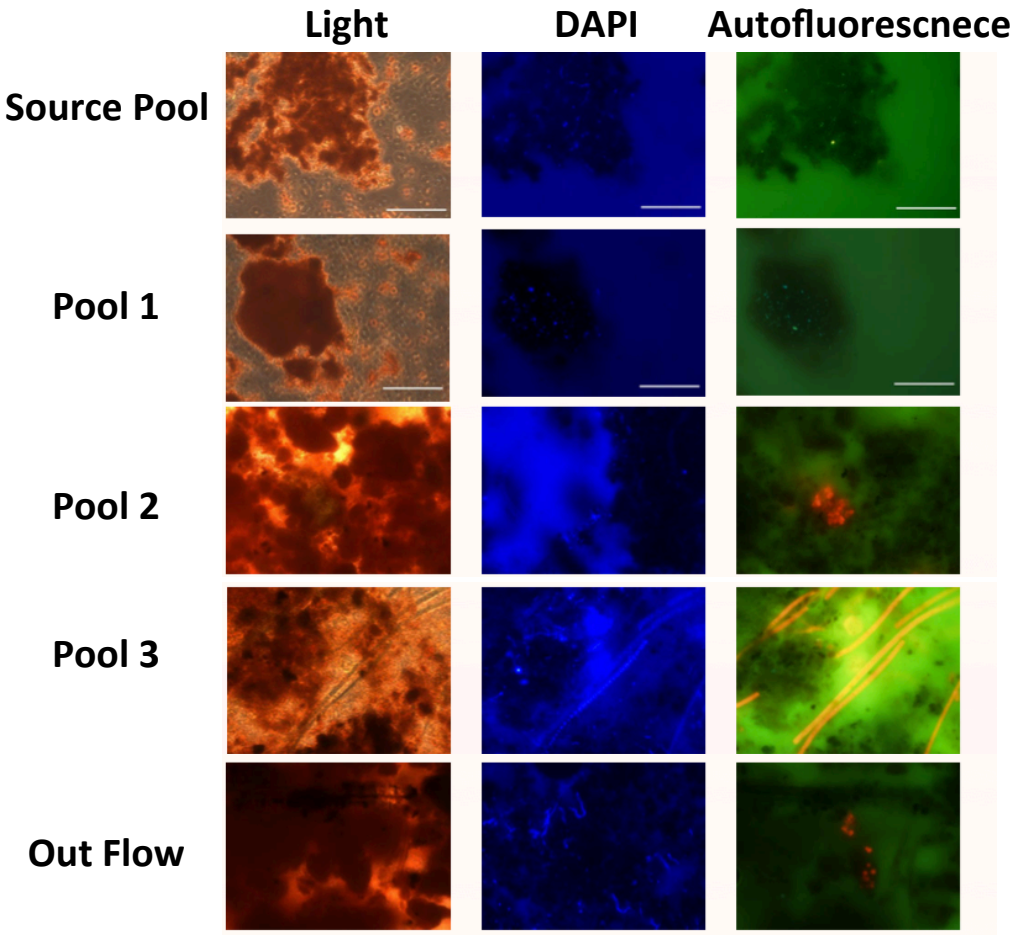


Figure 3

Microscopy images of sediment (Source and Pool 1) or mat (Pool 2, Pool 3, and Out Flow). Left are light microscopy images. Center and right are fluorescence images. At center, blue signal is DAPI-stained (Excitation: 365nm, Emission: BP445~50nm). At right, red is autofluorescence signal of Cyanobacteria (BP395~440nm, LP470nm). Scale bars 50 μm.

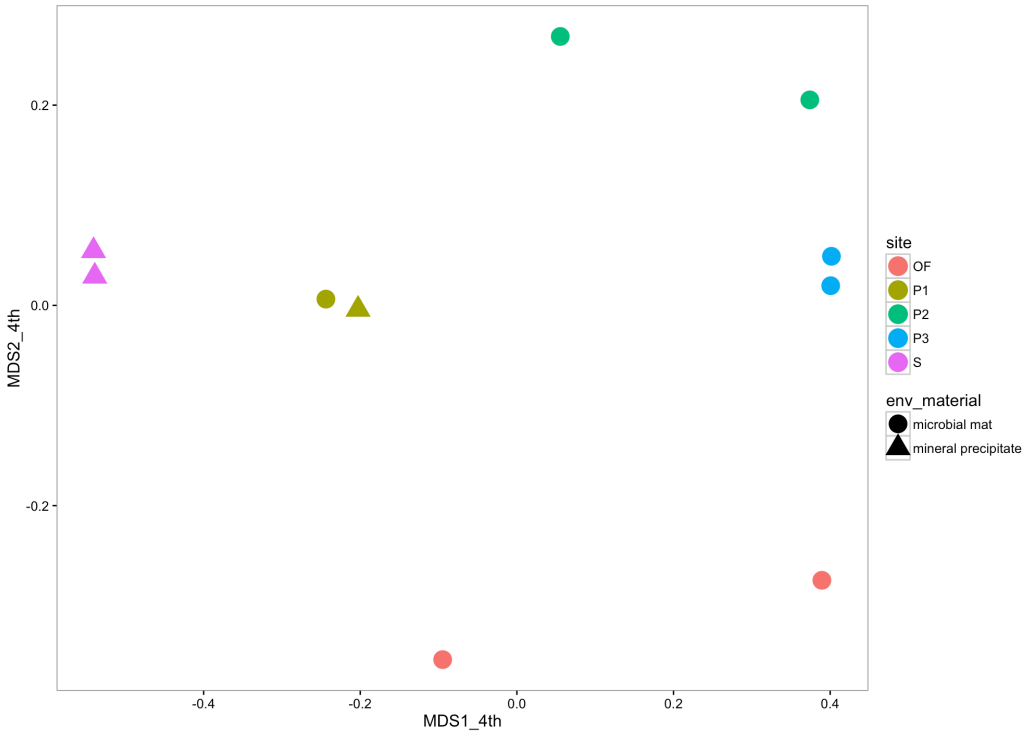


Figure 4:

Multidimensional scaling plot of Jinata samples. Each point represents the recovered microbial community from a given sample, with sites identified by color and sample type by shape. Samples plotting close to each other are relatively more similar in community composition. Abundance data are transformed by the 4th root to down-weight the effect of abundant taxa. Stress value is 0.0658.

T	63°C
pH	5.4
DO	4.7 μM
Fe^{2+}	261 μM
NH_3/NH_4^+	87 μM
Cl^-	654 mM
SO_4^{--}	17 mM
NO_3^-	b.d.
NO_2^-	b.d.
HPO_4^-	b.d.

Table 1:

Geochemical characteristics of Jinata source water. Nitrate, nitrite, and phosphate were below the detection limit of ~0.1 mg/L.

	pH	T (°C)	Fe(II) (µM)	DO (µM)	Descriptions
Source	**5.4**	**60-63**	**260**	**4.7 (source) 39 (surface)**	**Fluffy red iron oxide precipitate**
Pool 1	**5.8**	**59-59.5**	**265**	**58**	**Reddish precipitate and streamers in shallower regions, more yellowish deeper**
Pool 2	**6.5**	**44.5-54**	**151**	**134**	**Iron oxide-coated microbial mats. Orange to orange-green.**
Pool 3	**6.7**	**37.3-46**	**100**	**175**	**Green or mottled orange-green microbial mats, commonly with 1-5cm finger-like morphology.**
Outflow	**6.5**	**27-32**	**45**	**234**	**Ocean water within mixing zone at high tide, with constant flow of spring water from Pool 2. Thin green microbial mats.**

Table 2:
Summary table, showing overall geochemical transition from iron-rich hot spring source to dominantly marine outflow.

Sample:	Reads:	OTUs (99%)	Good Coverage (99%):	Shannon Index (99%):	Inverse Simpson (99%):	OTUs (97%)	Goods Coverage (97%)	Shannon Index (97%)	Inverse Simpson (97%):
Source bottom	26057	9558	0.724	10.594	83.020	4632	0.884	8.196	23.035
Source top	49340	14392	0.790	10.275	44.714	5530	0.932	7.229	12.835
Pool 1 A	97445	21166	0.848	10.128	56.287	10160	0.935	8.080	24.682
Pool 1 B	57250	10559	0.872	8.794	33.323	4766	0.945	6.414	12.005
Pool 2 A	41515	13114	0.759	9.754	24.340	7710	0.873	8.118	14.702
Pool 2 B	45171	17211	0.697	10.708	50.836	10525	0.832	8.980	25.783
Pool 3 A	45148	15988	0.722	10.287	33.295	9302	0.853	8.351	16.880
Pool 3 B	29778	12023	0.682	10.894	84.725	6625	0.837	8.553	31.520
Outflow A	32382	17741	0.542	11.931	57.572	11290	0.738	10.262	28.674
Outflow B	32651	8881	0.797	9.237	28.728	4210	0.909	6.373	9.850

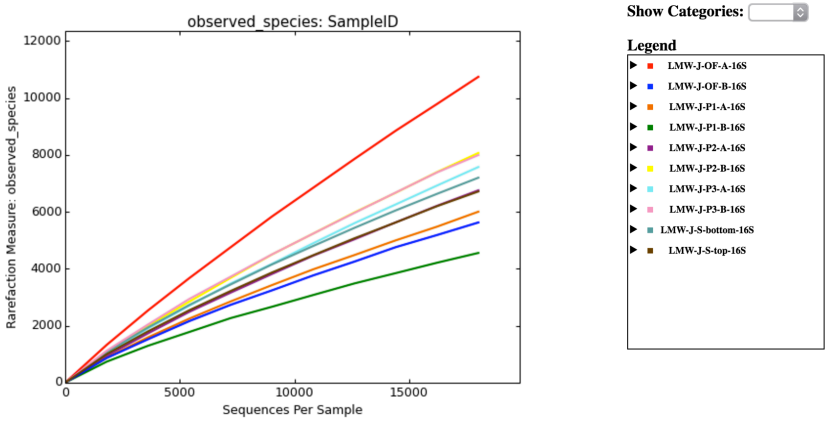
Table 3:
Diversity metrics of Jinata sequencing. Diversity metrics calculated for both 99% and 97% sequence identity cutoffs for assigning OTUs.

Taxon	S.A	S.B	P1.A	P1.B	P2.A	P2.B	P3.A	P3.B	OF.A	OF.B	
Bacteria;__Cyanobacteria;__Cyanobacteria;__SubsectionIII;__FamilyI	0.06%	0.02%	2.20%	0.48%	0.32%	17.63%	29.96%	23.25%	22.36%	37.44%	13.37%
Bacteria;__Proteobacteria;__Zetaproteobacteria;__Mariprofundales;__Mariprofundaceae	14.07%	29.96%	24.89%	29.72%	4.26%	1.89%	1.52%	0.78%	3.28%	0.54%	11.09%
Bacteria;__Aquificae;__Aquificae;__Aquificales;__Hydrogenothermaceae	30.54%	34.24%	3.54%	25.38%	0.21%	0.08%	0.16%	0.20%	0.55%	0.59%	9.55%
Bacteria;__Cyanobacteria;__Chloroplast;__o;__f	0.10%	0.01%	1.39%	0.15%	0.20%	0.14%	0.22%	0.40%	0.56%	33.11%	3.63%
Bacteria;__Chloroflexi;__Anaerolineae;__Anaerolineales;__Anaerolineaceae	5.66%	1.68%	4.45%	1.69%	3.23%	9.49%	2.87%	5.41%	0.30%	0.28%	3.50%
Bacteria;__Bacteroidetes;__Sphingobacteriia;__Sphingobacteriales;__Chitinophagaceae	0.05%	0.04%	0.08%	0.09%	0.15%	5.71%	10.93%	9.32%	1.18%	0.15%	2.77%
Bacteria;__Cyanobacteria;__Cyanobacteria;__SubsectionII;Other	0.00%	0.00%	0.07%	0.08%	24.47%	0.07%	0.19%	0.32%	0.10%	0.41%	2.57%
Bacteria;__Bacteroidetes;__Sphingobacteriia;__Sphingobacteriales;__Saprospiraceae	0.01%	0.02%	0.40%	0.12%	5.70%	6.53%	1.84%	0.34%	4.36%	3.02%	2.23%
Bacteria;__Planctomycetes;__Phycisphaerae;__Phycisphaerales;__Phycisphaeraceae	0.00%	0.00%	4.47%	0.08%	3.22%	2.75%	3.32%	6.68%	1.14%	0.15%	2.18%
Bacteria;__Cyanobacteria;__Cyanobacteria;Other;Other	0.01%	0.01%	0.18%	0.13%	0.21%	0.44%	3.21%	10.82%	0.71%	3.71%	1.94%
Bacteria;__Proteobacteria;Other;Other;Other	1.17%	1.92%	2.25%	10.48%	0.58%	0.50%	0.25%	0.22%	0.87%	0.08%	1.83%
Bacteria;__Acidobacteria;__Subgroup_22;__o;__f	0.53%	0.34%	5.21%	1.39%	8.16%	0.21%	0.36%	0.32%	0.18%	0.12%	1.68%
Bacteria;__Verrucomicrobia;__Opitutae;__Opitutales;__Opitutaceae	0.00%	0.00%	0.02%	0.03%	0.05%	2.23%	6.64%	1.16%	2.05%	0.05%	1.22%
Bacteria;__Proteobacteria;__Alphaproteobacteria;__Rhodobacterales;__Rhodobacteraceae	0.34%	0.26%	0.72%	0.31%	1.98%	2.00%	2.05%	0.92%	1.79%	1.19%	1.15%
Bacteria;__Proteobacteria;__Deltaproteobacteria;__Desulfuromonadales;Other	0.02%	0.02%	0.11%	0.16%	0.03%	0.68%	0.43%	0.21%	9.76%	0.07%	1.15%
Bacteria;__Proteobacteria;__Gammaproteobacteria;Other;Other	2.64%	1.74%	2.26%	1.99%	0.31%	0.55%	0.25%	0.26%	0.60%	0.25%	1.08%
Bacteria;__Deferribacteres;__Deferribacteres;__Deferribacterales;__Family_Incertae_Sedis	0.85%	0.20%	2.07%	0.40%	1.94%	0.32%	1.43%	0.57%	0.16%	0.09%	0.80%
Bacteria;__Planctomycetes;__Planctomycetacia;__Planctomycetales;__Planctomycetaceae	0.26%	0.06%	1.08%	0.60%	3.37%	1.06%	0.32%	0.17%	0.22%	0.17%	0.73%
Bacteria;__Chloroflexi;__Caldilineae;__Caldilineales;__Caldilineaceae	0.16%	0.06%	1.07%	0.20%	0.99%	0.64%	1.35%	0.44%	0.21%	0.02%	0.51%
Bacteria;__Nitrospirae;__Nitrospira;__Nitrospirales;__Nitrospiraceae	2.18%	0.30%	0.66%	0.26%	1.15%	0.10%	0.03%	0.05%	0.03%	0.07%	0.48%
Bacteria;__Proteobacteria;__Alphaproteobacteria;__Rhodospirillales;__Rhodospirillaceae	0.05%	0.05%	1.02%	0.81%	0.78%	0.71%	0.32%	0.11%	0.70%	0.09%	0.46%
Archaea;__Thaumarchaeota;__Marine_Group_I;Other;Other	0.00%	0.00%	0.01%	0.00%	3.09%	0.19%	0.01%	0.03%	0.02%	0.02%	0.34%
Bacteria;__Candidate_division_WS3;__c;__o;__f	0.15%	0.03%	1.21%	0.23%	0.15%	0.08%	0.07%	0.03%	0.12%	0.09%	0.22%
Bacteria;__Proteobacteria;__Gammaproteobacteria;__Chromatiales;__Chromatiaceae	0.23%	0.24%	0.21%	0.08%	0.23%	0.02%	0.03%	0.02%	0.03%	0.00%	0.11%
Bacteria;__Proteobacteria;__Alphaproteobacteria;__Rhodospirillales;Other	0.06%	0.01%	0.33%	0.09%	0.42%	0.01%	0.05%	0.03%	0.03%	0.00%	0.09%

Table 4:

Relative abundance of taxa to the Family level. Overall 10 most abundant taxa listed, as well as other taxa of interest mentioned in the text.

Supplemental Information:



Supplemental Figure 1:

Rarefaction curves of Jinata samples. Sampling depth is down sampled to 18000 reads for sample.

	S-top	OF-B	OF-A	S-bottom	P2-A	P2-B	P1-A	P3-A	P3-B	P1-B
S-top	0.000	0.934	0.772	0.219	0.809	0.843	0.553	0.829	0.870	0.268
OF-B		0.000	0.651	0.964	0.863	0.808	0.844	0.669	0.665	0.911
OF-A			0.000	0.783	0.797	0.500	0.662	0.461	0.551	0.734
S-bottom				0.000	0.797	0.830	0.550	0.829	0.866	0.320
P2-A					0.000	0.755	0.612	0.803	0.820	0.744
P2-B						0.000	0.692	0.442	0.480	0.790
P1-A							0.000	0.696	0.735	0.426
P3-A								0.000	0.268	0.795
P3-B									0.000	0.840
P1-B										0.000

Supplemental Table 1:

Weighted Unifrac matrix, showing the dissimilarity between samples, incorporating both differences in presence/absence as well as relative abundance of taxa (color coded to highlight relative dissimilarity—dark green is <0.33 dissimilarity, pale green >0.33 and <0.5, yellow >0.5 and <0.66, orange >0.66).

References

1. Anbar, a D. & Knoll, a H. Proterozoic ocean chemistry and evolution: a bioinorganic bridge? *Science* **297,** 1137–1142 (2002).
2. Baldani, José Ivo, et al. "The family rhodospirillaceae." *The Prokaryotes*. Springer Berlin Heidelberg, 2014. 533-618.
3. Barns, Susan M., Shannon L. Takala, and Cheryl R. Kuske. "Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment." *Applied and environmental microbiology* 65.4 (1999): 1731-1737.
4. Bird, L. J., Bonnefoy, V. & Newman, D. K. Bioenergetic challenges of microbial iron metabolisms. Trends Microbiol. 19, 330–340 (2011).
5. Bosak, T. et al. Cyanobacterial diversity and activity in modern conical microbialites. Geobiology 10, 384–401 (2012).
6. Brocks, J. J. et al. Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea. Nature 437, 866–870 (2005).
7. Brown, Igor I., et al. "Polyphasic characterization of a thermotolerant siderophilic filamentous cyanobacterium that produces intracellular iron deposits." *Applied and environmental microbiology* 76.19 (2010): 6664-6672.
8. Bushnell, B. "BBMap short read aligner." *University of California, Berkeley, California. URL http://sourceforge. net/projects/bbmap* (2016).
9. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., & Madden T.L. (2008) "BLAST+: architecture and applications." BMC Bioinformatics 10:421.

10. Canfield, D. E., Rosing, M. T. & Bjerrum, C. Early anaerobic metabolisms. *Philos. Trans. R. Soc. B Biol. Sci.* **361,** 1819–1836 (2006).

11. Caporaso, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6, 1621–1624 (2012).

12. Caporaso, J. Gregory, et al. "QIIME allows analysis of high-throughput community sequencing data." Nature methods 7.5 (2010): 335-336.

13. Case, R. J. *et al.* Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.* **73,** 278–288 (2007).

14. Catling, D. C., Glein, C. R., Zahnle, K. J. & McKay, C. P. Why O2 is required by complex life on habitable planets and the concept of planetary 'oxygenation time'. *Astrobiology* **8,** 377–395 (2005).

15. Chan, Clara S., et al. "Lithotrophic iron-oxidizing bacteria produce organic stalks to control mineral growth: implications for biosignature formation." *The ISME journal* 5.4 (2011): 717-727.

16. Chan, Clara S., et al. "The architecture of iron microbial mats reflects the adaptation of chemolithotrophic iron oxidation in freshwater and marine environments." *Frontiers in Microbiology* 7 (2016).

17. Chan, C. S., Emerson, D. & Luther, G. W. The role of microaerophilic Fe-oxidizing micro-organisms in producing banded iron formations. *Geobiology* **14,** 509–528 (2016b).

18. Crowe, S. A. et al. Photoferrotrophs thrive in an Archean Ocean analogue. Proc. Natl. Acad. Sci. 105, 15938–15943 (2008).

19. Daims, Holger. "The family nitrospiraceae." *The Prokaryotes*. Springer Berlin Heidelberg, 2014. 733-749.

20. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460-2461.

21. Ehrenreich, A., and Widdel, F., 1994, Anaerobic oxidation of ferrous iron by purple bacteria, a new type of phototrophic metabolism: Applied and Environmental Microbiology, v. 60, p. 4517–4526

22. Emerson, D. *et al*. A novel lineage of proteobacteria involved in formation of marine Fe-oxidizing microbial mat communities. *PLoS One* **2,** (2007).

23. Emerson, D., Fleming, E. J. & McBeth, J. M. Iron-oxidizing bacteria: an environmental and genomic perspective. Annu. Rev. Microbiol. 64, 561–583 (2010).

24. Fischer, W., Hemp, J. & Johnson, J. E. Evolution of Oxygenic Photosynthesis. Annu. Rev. Earth Planet. Sci. 44, (2016).

25. Fukunaga, Y. *et al*. Phycisphaera mikurensis gen. nov., sp nov., isolated from a marine alga, and proposal of Phycisphaeraceae fam. nov., Phycisphaerales ord. nov and Phycisphaerae classis nov in the phylum Planctomycetes. *J. Gen. Appl. Microbiol*. **55,** 267–275 (2009).

26. Garrity, George M., et al. "Phylum BIX. Deferribacteres phy. nov." *Bergey's Manual® of Systematic Bacteriology*. Springer New York, 2001. 465-471.

27. Geider, Richard J., et al. "Primary productivity of planet earth: biological determinants and physical constraints in terrestrial and aquatic habitats." *Global Change Biology* 7.8 (2001): 849-882.

28. GOOD, I.J., 1953, The population frequencies of species and the estimation of population parameters: Biometrika, v. 40, p. 237–264.

29. Götz, D., et al. "Persephonella marina gen. nov., sp. nov. and Persephonella guaymasensis sp. nov., two novel, thermophilic, hydrogen-oxidizing microaerophiles from deep-sea hydrothermal vents." International Journal of Systematic and Evolutionary Microbiology 52.4 (2002): 1349-1359.

30. Hemp J, Ward LM, Pace LA, Fischer WW. 2015. Draft genome sequence of Levilinea saccharolytica KIBI-1, a member of the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01357-15.

31. Hemp J, Ward LM, Pace LA, Fischer WW. 2015. Draft genome sequence of Ornatilinea apprima P3M-1, an anaerobic member of the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01353-15.

32. Hemp J, Ward LM, Pace LA, Fischer WW. 2015. Draft genome sequence of Ardenticatena maritima 110S, a thermophilic nitrate- and iron-reducing member of the Chloroflexi class Ardenticatenia. Genome Announc 3(6):e01347-15.

33. HILL, M.O., 1973, Diversity and evenness: a unifying notation and its consequences: Ecology, v. 54, p. 427–432.

34. Hirayama, Hisako, et al. "Methylomarinovum caldicuralii gen. nov., sp. nov., a moderately thermophilic methanotroph isolated from a shallow submarine hydrothermal system, and proposal of the family Methylothermaceae fam. nov." *International journal of systematic and evolutionary microbiology* 64.3 (2014): 989-999.

35. Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press.

36. Hurowitz, Joel A., et al. "Origin of acidic surface waters and the evolution of atmospheric chemistry on early Mars." *Nature Geoscience* 3.5 (2010): 323-326.

37. Imhoff, Johannes F. "The Family Chromatiaceae." *The Prokaryotes*. Springer Berlin Heidelberg, 2014. 151-178.

38. Johnson, J. E., Gerpheide, A., Lamb, M. P. & Fischer, W. W. O2 constraints from Paleoproterozoic detrital pyrite and uraninite. *Geol. Soc. Am. Bull.* **126,** 813–830 (2014).

39. Johnston, David T., et al. "Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age." *Proceedings of the National Academy of Sciences* 106.40 (2009): 16925-16929.

40. Kale, V. et al. Litorilinea aerophila gen. nov., sp. nov., an aerobic member of the class Caldilineae, phylum Chloroflexi, isolated from an intertidal hot spring. Int. J. Syst. Evol. Microbiol. 63, 1149–1154 (2013).

41. Kaneoka, Ichiro, Naoki Isshiki, and Shigeo Zashu. "K-Ar ages of the Izu-Bonin islands." *Geochemical Journal* 4.2 (1970): 53-60.

42. Kasting, James F. "What caused the rise of atmospheric O 2?." *Chemical Geology* 362 (2013): 13-25.

43. Kato, S. *et al*. Comparative genomic insights into ecophysiology of neutrophilic, microaerophilic iron oxidizing bacteria. *Front. Microbiol*. **6,** 1–16 (2015).

44. Katoh, Kazutaka, et al. "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform." Nucleic acids research 30.14 (2002): 3059-3066.

45. Kharecha, P., Kasting, J. & Siefert, J. A coupled atmosphere–ecosystem model of the early Archean Earth. *Geobiology* **3,** 53–76 (2005).

46. Kielak, Anna M., et al. "The ecology of Acidobacteria: moving beyond genes and genomes." *Frontiers in Microbiology* 7 (2016).

47. Klatt, C. G. et al. Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. ISME J. 5, 1262–1278 (2011).

48. Klueglein, N., and A. Kappler. Abiotic oxidation of Fe (II) by reactive nitrogen species in cultures of the nitrate-reducng Fe(II) oxidizer Acidovorax sp. BoFeN1—questioning the existence of enzymatic Fe(II) oxidation. Geobiology 11.2 (2013): 180-190.

49. Konhauser, K. O. *et al*. Could bacteria have formed the Precambrian banded iron formations? *Geology* **30,** 1079–1082 (2002).

50. Krepski, S. T., Emerson, D., Hredzak-Showalter, P. L., Luther III, G. W. & Chan, C. S. Morphology of biogenic iron oxides records microbial physiology and environmental conditions: toward interpreting iron microfossils. *Geobiology* **11,** 457–471 (2013).

51. Lebedeva EV, Off S, Zumbragel S, Kruse M, Shagzhina A, Lucker S, Maixner F, Lipski A, Daims H, Spieck E (2011) Isolation and characterization of a moderately thermophilic nitrite-oxidizing bacterium from a geothermal spring. FEMS Microbiol

Ecol 75:195–204

52. Li H.*, Handsaker B.*, Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics, 25, 2078-9. [PMID: 19505943]

53. LOZUPONE, C., AND KNIGHT, R., 2005, UniFrac: a new phylogenetic method for comparing microbial communities: Applied and Environmental Microbiology, v. 71, p. 8228–8235.

54. McIlroy, Simon Jon, and Per Halkjær Nielsen. "The family saprospiraceae." *The Prokaryotes*. Springer Berlin Heidelberg, 2014. 863-889.

55. Meyer-Dombard, D., Jan P. Amend, and Magdalena R. Osburn. "Microbial diversity and potential for arsenic and iron biogeochemical cycling at an arsenic rich, shallow-sea hydrothermal vent (Tutum Bay, Papua New Guinea)." Chemical Geology 348 (2013): 37-47.

56. Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010) "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1 - 8.

57. Neubauer, S. C. et al. Life at the energetic edge: kinetics of circumneutral Fe oxidation by lithotrophic iron oxidizing bacteria isolated from the wetland plant rhizosphere. Appl. Environ. Microbiol. 68, 3988–3995 (2002).

58. Oksanen, Jar, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens and Helene Wagner (2016). vegan: Community Ecology Package. R package version 2.3-5. https://CRAN.R-project.org/package=vegan

59. Pace LA, Hemp J, Ward LM, Fischer WW. 2015. Draft genome of Thermanaerothrix daxensis GNS-1, a thermophilic facultative anaerobe from the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01354-15.

60. PARADA, A., NEEDHAM, D.M., AND FUHRMAN, J.A., 2015, Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time-series and global field samples: Environmental Microbiology, v. 18, p. 1403–1414.

61. Pester, Michael, Christa Schleper, and Michael Wagner. "The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology." *Current opinion in microbiology* 14.3 (2011): 300-306.

62. Pierson, B. K., Paranteau, M. N. & Griffin, B. M. Phototrophs in high-iron-concentration microbialmats: Physiological ecology of phototrophs in an iron-depositing hot spring. Appl. Environ.Microbiol. 65, 5474–5483 (1999). 36.

63. Poulton, Simon W., Philip W. Fralick, and Donald E. Canfield. "Spatial variability in oceanic redox structure 1.8 billion years ago." *Nature Geoscience* 3.7 (2010): 486-490.

64. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments. Plos One 5(3).

65. Pujalte, María J., et al. "The family Rhodobacteraceae." *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* (2014): 439-512.

66. Quast, Christian, et al. "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools." Nucleic acids research41.D1 (2013): D590-D596.

67. R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

68. Rantamäki, S. et al. Oxygen produced by cyanobacteria in simulated Archaean conditions partly oxidizes ferrous iron but mostly escapes—conclusions about early evolution. Photosynth. Res. 1–9 (2016).

69. Rodrigues, Jorge LM, and Jantiya Isanapong. "The family Opitutaceae." *The Prokaryotes*. Springer Berlin Heidelberg, 2014. 751-756.

70. Roeselers, G. et al. Diversity of phototrophic bacteria in microbial mats from Arctic hot springs (Greenland). Environ. Microbiol. 9, 26–38 (2007).

71. Rosenberg, Eugene. "The family Chitinophagaceae." *The Prokaryotes*. Springer Berlin Heidelberg, 2014. 493-495.

72. Sahoo SK, et al. (2012) Ocean oxygenation in the wake of the Marinoan glaciation. Nature 489(7417):546–549.

73. Scott, C., et al. "Tracing the stepwise oxygenation of the Proterozoic ocean." *Nature* 452.7186 (2008): 456-459.

74. Sekiguchi, Y. et al. Anaerolinea thermophila gen. nov., sp. nov. and Caldilinea aerophila gen. nov., sp. nov., novel filamentous thermophiles that represent a

previously uncultured lineage of the domain Bacteria at the subphylum level. Int. J. Syst. Evol. Microbiol. 53, 1843–1851 (2003).

75. Shannon, C. E. (1948) A mathematical theory of communication. The Bell System Technical Journal, 27, 379–423 and 623–656.

76. Shih, P. M. & Matzke, N. J. Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 12355–60 (2013).

77. Shih, P., Hemp, J., Ward, L., Matzke, N. & Fischer, W. 2016. Crown group oxyphotobacteria postdate the rise of oxygen. Geobiology.

78. SIMPSON, E.H., 1949, Measurement of diversity: Nature, v. 163, p. 688.

79. Skennerton, C. T. *et al*. Genomic reconstruction of an uncultured hydrothermal vent gammaproteobacterial methanotroph (family Methylothermaceae) indicates multiple adaptations to oxygen limitation. *Front. Microbiol.* **6,** 1–12 (2015).

80. A. Stamatakis: "RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies". In Bioinformatics, 2014

81. Stamenković V, LM Ward, and WW Fischer. 2017. Supercool oxygen Oases on Mars. In prep.

82. Stookey, Lawrence L. "Ferrozine---a new spectrophotometric reagent for iron." Analytical chemistry 42.7 (1970): 779-781.

83. Swanner, E. D. et al. Modulation of oxygen production in Archaean oceans by episodes of Fe(II) toxicity. Nat. Geosci. 8, 126–130 (2015).

84. Takai, Ken, and Satoshi Nakagawa. "The Family Hydrogenothermaceae."The Prokaryotes. Springer Berlin Heidelberg, 2014. 689-699.

85. Trembath-Reichert E, Ward LM, Slotznick SP, Bachtel SL, Kerans C, Grotzinger JP, Fischer WW (2016) Gene sequencing microbial community analysis of mat morphologies, Caicos Platform, British West Indies, Journal of Sedimentary Research.

86. Walker, J. C. & Brimblecombe, P. Iron and sulfur in the pre-biologic ocean. *Precambrian Res*. **28,** 205–222 (1985).

87. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microb 73(16): 5261-5267.

88. Ward LM, Hemp J, Pace LA, Fischer WW. 2015. Draft genome sequence of Leptolinea tardivitalis YMTK-2, a mesophilic anaerobe from the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01356-15.

89. Ward, L. M., Kirschvink, J. L. & Fischer, W. W. 2016. Timescales of Oxygenation Following the Evolution of Oxygenic Photosynthesis. Orig. Life Evol. Biosph. 46(1) pp51-65.

90. Ward LM, A Idai, T Kakegawa, WW Fischer, and SE McGlynn. 2017. Microbial diversity and iron oxidation at Okuoku-hachikurou Onsen, a Japanese hot spring analog of Precambrian iron formation. Geobiology, in revision.

91. Waterhouse, A.M., Procter, J.B., Martin, D.M.A, Clamp, M. and Barton, G. J. (2009) "Jalview Version 2 - a multiple sequence alignment editor and analysis workbench" Bioinformatics25 (9) 1189-1191 doi: 10.1093/bioinformatics/btp033

92. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009.

93. Yamada, T. et al. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the. Int. J. Syst. Evol. Microbiol. 56, 1331–1340 (2006).

94. Yoon, J., Jang, J. H. & Kasai, H. Algisphaera agarilytica gen nov, sp nov, a novel representative of the class Phycisphaerae within the phylum Planctomycetes isolated from a marine alga. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **105,** 317–324 (2014)

95. Youssef, N. H. *et al*. In Silico Analysis of the Metabolic Potential and Niche Specialization of Candidate Phylum 'Latescibacteria' (WS3). *PLoS One* **10,** e0127499 (2015).

96. Zerkle, A. L. *et al*. Onset of the aerobic nitrogen cycle during the Great Oxidation Event. *Nature* 1–10 (2017). doi:10.1038/nature20826

A p p e n d i x 4

GENOME SEQUENCING OF DIVERSE CHLOROFLEXI

Draft genome of *Leptolinea tardivitalis* YMTK-2, a mesophilic anaerobe from the Chloroflexi class *Anaerolineae*

Ward LM, Hemp J, Pace LA, Fischer WW. 2015. Draft genome sequence of Leptolinea tardivitalis YMTK-2, a mesophilic anaerobe from the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01356-15. DOI: 10.1128/genomeA.01356-15

Abstract

We present the draft genome of *Leptolinea tardivitalis* YMTK-2, a member of the Chloroflexi phylum. This organism was initially characterized as a strictly anaerobic non-motile fermenter; however, genome analysis demonstrates that it encodes for a flagella and might be capable of aerobic respiration.

Genome announcement

Leptolinea tardivitalis YMTK-2 was originally isolated from sludge granules of an upflow anaerobic sludge blanket (UASB) reactor used in wastewater processing (1). Closely related stains have been reported from other anaerobic wastewater treatment systems, soils (2), and aquatic moss pillars from Antarctica (3). *L. tardivitalis* is a filamentous, non-

sporulating organism that can ferment a number of sugars and fatty acids (1). It grows optimally at 37 °C (range 25-50 °C) and pH 7.0 (range pH 6.0-7.2).

The genome of *Leptolinea tardivitalis* YMTK-2 (DSM 16556) was sequenced as part of a project to expand the phylogenetic breadth of Chloroflexi genomes. Genome sequencing was performed at Seqmatic using the Illumina MiSeq sequencing platform. SPAdes 3.1.1 (4) was used to assemble the genome. The genome was screened for contaminants based on sequence coverage, GC composition, and BLAST hits of conserved single copy genes. Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The draft genome is 3.74 Mb in size, assembled into 15 contigs. It encodes 3349 genes, 2939 CDSs, 1 16sRNA, 46 tRNAs, and 5 CRISPR arrays. It is estimated to be ~96% complete based on conserved single copy genes (107/111).

All *Anaerolineae* strains isolated to date have been classified as strictly anaerobic fermenters (1). However genome analysis suggests that *L. tardivitalis* has a richer physiology than previously recognized. It encodes for both Complex I (NADH dehydrogenase) and quinol *bd* oxidase (5), suggesting that it might be capable of microaerobic respiration. *L. tardivitalis* is missing genes for LPS biosynthesis and known outer membrane proteins, suggesting that it does not have an outer membrane (6). Furthermore *L. tardivitalis* encodes for gram-positive flagella, making it likely that it is motile under certain conditions.

References:

1. **Yamada T**, **Sekiguchi Y**, **Hanada S**, **Imachi H**, **Ohashi A**, **Harada H**, **Kamagata Y**. 2006. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. Int J Syst Evol Microbiol **56**:1331–1340.

2. **Behrens S**, **Azizian MF**, **McMurdie PJ**, **Sabalowsky A**, **Dolan ME**, **Semprini L**, **Spormann AM**. 2008. Monitoring abundance and expression of "Dehalococcoides" species chloroethene-reductive dehalogenases in a tetrachloroethene-dechlorinating flow column. Appl Environ Microbiol **74**:5695–5703.

3. **Nakai R**, **Abe T**, **Baba T**, **Imura S**, **Kagoshima H**, **Kanda H**, **Kanekiyo A**, **Kohara Y**, **Koi A**, **Nakamura K**, **Narita T**, **Niki H**, **Yanagihara K**, **Naganuma T**. 2011. Microflorae of aquatic moss pillars in a freshwater lake, East Antarctica, based on fatty acid and 16S rRNA gene analyses. Polar Biol **35**:425–433.

4. **Nurk S**, **et al**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20**:714–737.

5. **Borisov VB**, **Gennis RB**, **Hemp J**, **Verkhovsky MI**. 2011. The cytochrome bd respiratory oxygen reductases. Biochim Biophys Acta **1807**:1398–1413.

6. **Sutcliffe IC**. 2011. Cell envelope architecture in the Chloroflexi: a shifting frontline in a phylogenetic turf war. Environ Microbiol **13**:279–282.

Draft genome of *Bellilinea caldifistulae* reveals capacity for aerobic respiration and phototrophy in the *Chloroflexi* class *Anaerolinea*

Abstract

We report the draft genome of *Bellilinea caldifistulae*, a member of the *Anaerolineae* class of the bacterial phylum *Chloroflexi*. This genome contains genes for aerobic respiration and proteorhodopsin-based phototrophy that were not revealed by culture-based studies, and expands the known metabolic potential of *Anaerolineae.*

Genome Announcement

The bacterial phylum *Chloroflexi* (previously referred to as the Green Non-sulfur Bacteria) is characterized by filamentous organisms with gliding motility; its studied members are dominantly anoxygenic photoheterotrophs (1). The majority of cultured Chloroflexi belong to the class *Chloroflexi*, which includes many anoxygenic phototrophs and facultative aerobes. In contrast, the class *Anaerolineae* is less well characterized and only contains members described as anaerobic fermenters (1,2). But this group makes up more than 70% of all *Chloroflexi* sequences present in 16S gene datasets (3).

We report here the draft genome sequence for *Bellilinea caldifistulae,* isolated from a propionate-degrading consortium of thermophilic digester sludge in Japan, and described in pure culture (4). *B. caldifistulae* was characterized as obligately anaerobic, nonmotile

filamentous, and capable of growth on a range of carbohydrates when supplemented with yeast extract under circumneutral pH and temperatures between 45°C and 65°C (4).

The genome of *Bellilinea caldifistulae* was sequenced as part of a project to expand the phylogenetic breadth of Chloroflexi genomes and reconstruct their metabolic evolution. Genome sequencing was performed at Seqmatic using the Illumina MiSeq sequencing platform. SPAdes 3.1.1 (5) was used to assemble the genome. The genome was screened for contaminants based on sequence coverage, GC composition, and BLAST hits of conserved single copy genes. Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The draft genome is 4.41 Mb in size, assembled into 45 contigs. It encodes 3846 genes, 3347 CDSs, 2 16sRNA, 50 tRNAs, and 3 CRISPR arrays. It is estimated to be ~96% complete based on conserved single copy genes (107/111).

Interestingly, analysis of the annotated *B. caldifistulae* genome revealed genes encoding for traits not observed in culture studies. This includes the capacity for aerobic respiration marked by a complete high-potential electron transport chain (using alternative complex III), and terminal cytochrome *c* oxidase (A-family). The genome also contains a *bd* oxidase. This study continues the first description of the capacity for aerobic respiration in *Anaerolineae* (6), a class previously defined by its obligate anaerobic lifestyle (3). The genome also revealed the capacity for proteorhodopsin-based phototrophy; however no genes associated with phototrophic chlorin-based reaction centers were observed. Furthermore the *B. caldifistulae* genome contains genes for flagellar motility and chemotaxis, in contrast to its description as nonmotile.

This draft genome helps expand the genomic coverage of *Anaerolineae*—an understudied class within the *Chloroflexi*. Additionally, it reveals the potential capacity for metabolic traits common in other *Chloroflexi* that were previously undiscovered in *Anaerolineae,* providing valuable new data for phylogenetic investigations into the evolution of high-potential metabolisms in this phylum.

References

1. **Overmann J**. 2008. Green Nonsulfur Bacteria. *In* Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd, Chichester, UK.
2. **Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y**. 2006. *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacterial phylum *Chloroflexi*. Int J Syst Evol Microbiol **56**:1331-1340.
3. **Yamada T, Sekiguchi Y**. 2009. Cultivation of uncultured *Chloroflexi* subphyla: significance and ecophysiology of formerly uncultured *Chloroflexi* 'Subphylum I' with natural and biotechnological relevance. Microbes Environ. **24**(3):205-216.
4. **Yamada T, Imachi H, Ohashi A, Harada H, Hanada S, Kamagata Y, Sekiguchi Y**. 2007. *Bellilinea caldifistulae* gen. nov., sp. nov. and *Longilinea arvoryzae* gen. nov., sp. nov., strictly anaerobic, filamentous bacteria of the phylum *Chloroflexi* isolated from methanogenic propionate-degrading consortia. Int J Syst Evol Microbiol **57**:2299-2306.

5. **Nurk S**, **Bankevich A**, **Antipov D**, **Gurevich AA**, **Korobeynikov A**, **Lapidus A**, **Prjibelski AD**, **Pyshkin A**, **Sirotkin A**, **Sirotkin Y**, **Stepanauskas R**, **Clingenpeel SR**, **Woyke T**, **McLean JS**, **Lasken R**, **Tesler G**, **Alekseyev MA**, **Pevzner PA**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20**:714–737.

6. **Ward LM, Hemp J, Pace LA, Fischer WW**. 2015. Draft genome of Leptolinea tardivitalis YMTK-2, a mesophilic anaerobe from the Chloroflexi class Anaerolineae, Genome Announc, in press.

Draft genome of *Herpetosiphon geysericola* GC-42, a non-phototrophic member of the Chloroflexi class *Chloroflexi*a

Ward LM, Hemp J, Pace LA, Fischer WW. 2015. Draft genome sequence of Herpetosiphon geysericola GC-42, a nonphototrophic member of the Chloroflexi class Chloroflexia. Genome Announc 3(6):e01352-15. DOI: 10.1128/genomeA.01352-15

Abstract

We report the draft genome of *Herpetosiphon geysericola* GC-42, a predatory non-phototrophic member of the class *Chloroflexia* in the Chloroflexi phylum. This genome provides insight into the evolution of phototrophy and aerobic respiration within the Chloroflexi.

Genome announcement

The majority of cultured members of the bacterial phylum Chloroflexi belong to the class *Chloroflexia*, which are prominently anoxygenic phototrophs (1). However the most basal *Chloroflexia,* members of the orders *Herpetosiphonales* and *Kallotenuales,* are non-phototrophic (2). This makes these clades of central importance to understanding the evolution of phototrophy within the Chloroflexi. The *Herpetosiphonales* order currently contains only two species, *Herpetosiphon geysericola* and *Herpetosiphon aurantiacus,* which are pigmented filamentous organoheterotrophs that exhibit gliding motility (3). In culture *Herpetosiphon* strains are obligate aerobes that prefer microaerobic conditions (3). The physiology and ecology of these organisms is not well understood, but it has been

suggested that members of the *Herpetosiphon* genus might be capable of predation via a "wolf pack" strategy (4). *H. geysericola* was isolated from a biofilm at a hot spring in Baja California, Mexico (5).

The genome of *Herpetosiphon geysericola* GC-42 (DSM 7119) was sequenced as part of a project to expand the phylogenetic breadth of Chloroflexi genomes. Genome sequencing was performed at Seqmatic using the Illumina MiSeq sequencing platform. SPAdes 3.1.1 (6) was used to assemble the genome. The genome was screened for contaminants based on sequence coverage, GC composition, and BLAST hits of conserved single copy genes. Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The draft genome is 6.24 Mb in size, assembled into 46 contigs. It encodes 5335 genes, 4688 CDSs, 2 16sRNA, 47 tRNAs, and 6 CRISPR arrays. It is estimated to be ~99% complete based on conserved single copy genes (111/111).

Analysis of the *H. geysericola* genome revealed genes for a branched aerobic respiratory chain including two different Complex Is (NADH dehydrogenase), Complex II (succinate dehydrogenase), Complex III (cytochrome *bc* complex), an A-family heme-copper oxygen reductase, and a quinol *bd* oxidase. No genes for phototrophy were found. In addition no genes associated with the 3-hydroxypropionate cycle—a CO_2 fixation pathway present in photosynthetic members of Chloroflexi (7)—were found in *H. geysericola*; this is consistent with its predicted heterotrophic ecology.

The sequencing of *H. geysericola*, along with *Herpetosiphon aurantiacus* (4), completes the genomic knowledge of the cultured members of the genus *Herpetosiphon*. These data provide important constraints that help ordinate the aquisition of phototrophy, aerobic respiration, and carbon fixation within the Chloroflexi.

References

1. **Overmann J**. 2008. Green Nonsulfur Bacteria. *In* Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd, Chichester, UK.

2. **Cole JK**, **Gieler BA**, **Heisler DL**, **Palisoc MM**, **Williams AJ**, **Dohnalkova AC**, **Ming H**, **Yu TT**, **Dodsworth JA**, **Li W-J**, **Hedlund BP**. 2013. Kallotenue papyrolyticum gen. nov., sp. nov., a cellulolytic and filamentous thermophile that represents a novel lineage (Kallotenuales ord. nov., Kallotenuaceae fam. nov.) within the class Chloroflexia. Int J Syst Evol Microbiol **63**:4675–4682.

3. **Lee N**, **Reichenbach H**. 2006. The Genus Herpetosiphon, pp. 854–877. *In*. The Prokaryotes.

4. **Kiss H**, **Nett M**, **Domin N**, **Martin K**, **Maresca JA**, **Copeland A**, **Lapidus A**, **Lucas S**, **Berry KW**, **Glavina Del Rio T**, **Dalin E**, **Tice H**, **Pitluck S**, **Richardson P**, **Bruce D**, **Goodwin L**, **Han C**, **Detter JC**, **Schmutz J**, **Brettin T**, **Land M**, **Hauser L**, **Kyrpides NC**, **Ivanova N**, **Göker M**, **Woyke T**, **Klenk H-P**, **Bryant DA**. 2011. Complete genome sequence of the filamentous gliding predatory bacterium Herpetosiphon aurantiacus type strain (114-95(T)). Stand Genomic Sci **5**:356–370.

5. **Lewin RA**. 1970. New Herpetosiphon species (Flexibacterales). Can J Microbiol **16**:517–520.

6. **Nurk S**, **Bankevich A**, **Antipov D**, **Gurevich AA**, **Korobeynikov A**, **Lapidus A**, **Prjibelski AD**, **Pyshkin A**, **Sirotkin A**, **Sirotkin Y**, **Stepanauskas R**, **Clingenpeel SR**, **Woyke T**, **McLean JS**, **Lasken R**, **Tesler G**, **Alekseyev MA**, **Pevzner PA**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20**:714–737.

7. **Strauss G**, **Fuchs G**. 1993. Enzymes of a novel autotrophic CO2 fixation pathway in the phototrophic bacterium Chloroflexus aurantiacus, the 3-hydroxypropionate cycle. Eur J Biochem **215**:633–643.

Draft genome of *Levilinea saccharolytica* KIBI-1, a member of the Chloroflexi class *Anaerolineae*

Hemp J, Ward LM, Pace LA, Fischer WW. 2015. Draft genome sequence of Levilinea saccharolytica KIBI-1, a member of the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01357-15. DOI: 10.1128/genomeA.01357-15

Abstract

We report the draft genome of *Levilinea saccharolytica* KIBI-1, a facultative anaerobic member of the Chloroflexi class *Anaerolineae*. While *L. saccharolytica* was characterized as an obligate anaerobe, genome analysis provides evidence for the presence of both aerobic respiration and partial denitrification pathways.

Genome Announcement

Levilinea saccharolytica KIBI-1 was isolated from sludge granules of a mesophilic wastewater reactor (1). A closely related strain was detected in a trichlorobenzene-transforming microbial consortium (2). *L. saccharolytica* was characterized as an obligately anaerobic, non-motile filamentous microbe capable of growth on a range of carbohydrates when supplemented with yeast extract (1). It grows optimally at 37 °C and pH 7.0.

The genome of *Levilinea saccharolytica* KIBI-1 (DSM 16555) was sequenced as part of a project to expand the phylogenetic breadth of Chloroflexi genomes. Genome sequencing was performed at Seqmatic using the Illumina MiSeq sequencing platform. SPAdes 3.1.1

(3) was used to assemble the genome. The genome was screened for contaminants based on sequence coverage, GC composition, and BLAST hits of conserved single copy genes. Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The draft genome is 4.30 Mb in size, assembled into 65 contigs. It encodes 3672 genes, 3173 CDSs, 1 16sRNA, and 46 tRNAs. It is estimated to be ~99% complete based on conserved single copy genes (110/111).

The *Anaerolineae* described to date have all been classified as strict anaerobes, however *L. saccharolytica* encodes for a branched aerobic respiration pathway. It has a Complex I (NADH dehydrogenase), Complex II (succinate dehydrogenase), an Alternative Complex III (ACIII) (4), and both an A-family heme-copper oxygen reductase and a *bd* oxidase. It also encodes for two nitrite reduction pathways; a NirS nitrite reductase that reduces nitrite into nitric oxide, and a NrfA protein that reduces it into ammonia. The presence of aerobic respiration genes in *L. saccharolytica* and other recently sequenced *Anaerolineae* suggests that this Chloroflexi class is substantially more physiologically diverse than previously recognized (5).

References

1. **Yamada T**, **Sekiguchi Y**, **Hanada S**, **Imachi H**, **Ohashi A**, **Harada H**, **Kamagata Y**. 2006. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. Int J Syst Evol Microbiol **56**:1331–1340.

2. **Wintzingerode von F**, **Selent B**, **Hegemann W**, **Göbel UB**. 1999. Phylogenetic analysis of an anaerobic, trichlorobenzene-transforming microbial consortium. Appl

Environ Microbiol **65**:283–286.

3. **Nurk S**, **Bankevich A**, **Antipov D**, **Gurevich AA**, **Korobeynikov A**, **Lapidus A**, **Prjibelski AD**, **Pyshkin A**, **Sirotkin A**, **Sirotkin Y**, **Stepanauskas R**, **Clingenpeel SR**, **Woyke T**, **McLean JS**, **Lasken R**, **Tesler G**, **Alekseyev MA**, **Pevzner PA**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20**:714–737.

4. **Refojo PN**, **Ribeiro MA**, **Calisto F**, **Teixeira M**, **Pereira MM**. 2013. Structural composition of alternative complex III: Variations on the same theme. Biochim Biophys Acta **1827**:1378–1382.

5. **Hug LA**, **Castelle CJ**, **Wrighton KC**, **Thomas BC**, **Sharon I**, **Frischkorn KR**, **Williams KH**, **Tringe SG**, **Banfield JF**. 2013. Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. Microbiome **1**:22.

Draft genome of *Thermanaerothrix daxensis* GNS-1, a thermophilic facultative anaerobe from the Chloroflexi class *Anaerolineae*

Pace LA, Hemp J, Ward LM, Fischer WW. 2015. Draft genome of Thermanaerothrix daxensis GNS-1, a thermophilic facultative anaerobe from the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01354-15. DOI: 10.1128/genomeA.01354-15

Abstract

We present the draft genome of *Thermanaerothrix daxensis* GNS-1, a thermophilic member of the Chloroflexi phylum. This organism was initially characterized as a non-motile strictly anaerobic fermenter; however, genome analysis demonstrates that it encodes genes for a flagellum and multiple pathways for aerobic and anaerobic respiration.

Genome Announcement

Thermanaerothrix daxensis GNS-1 was isolated from a deep groundwater aquifer (149 m) housed within sedimentary strata of the large Mesozoic and Tertiary Aquitaine Basin in southwestern France (1). Closely related strains have been reported from a hot spring in Yellowstone National Park (2), a hot spring in southwestern Taiwan, geothermal soil, a thermophilic anaerobic digestive sludge, and a thermophilic electrochemical cell (3). *T. daxensis* is a filamentous, non-sporulating organism that can ferment a number of sugars and organic acids (1). It grows optimally at 65 °C (range 50-73 °C) and pH 7 (range pH 5.8-8.5) (1).

The genome of *Thermanaerothrix daxensis* GNS-1 (DSM 23592) was sequenced as part of a project to expand the phylogenetic breadth of Chloroflexi genomes. Genome sequencing was performed at Seqmatic using the Illumina MiSeq sequencing platform. SPAdes 3.1.1 (4) was used to assemble the genome. The genome was screened for contaminants based on sequence coverage, GC composition, and BLAST hits of conserved single copy genes. Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The draft genome is 3.06 Mb in size, assembled into 6 contigs. It encodes 2798 genes, 2395 CDSs, 1 16sRNA, 47 tRNAs, and 4 CRISPR arrays. It is estimated to be ~95% complete based on conserved single copy genes (106/111).

Genome analysis of *T. daxensis* detected the presence of aerobic and anaerobic respiration pathways, hinting at a richer physiology than previously recognized. It encodes for Complex I (NADH dehydrogenase), Complex II (succinate dehydrogenase), and an aerobic CO dehydrogenase. It also has two aerobic respiration modules; an A-family heme-copper oxygen reductase coupled to an alternative complex III (ACIII) (5), and a quinol *bd* oxidase (6). In addition *T. daxensis* has two respiratory nitrite reductases; NirS, which reduces NO_2^- to NO, and NrfA that reduces NO_2^- to NH_4^+. The genome provides no evidence for the presence of LPS biosynthesis genes or outer membrane proteins, suggesting that this organism has only one membrane (7). Furthermore it encodes for a gram-positive flagella and is likely motile under certain physiological conditions.

References

1. **Grégoire P**, **Fardeau M-L**, **Joseph M**, **Guasco S**, **Hamaide F**, **Biasutti S**, **Michotey V**, **Bonin P**, **Ollivier B**. 2011. Isolation and characterization of Thermanaerothrix daxensis gen. nov., sp. nov., a thermophilic anaerobic bacterium pertaining to the phylum "Chloroflexi," isolated from a deep hot aquifer in the Aquitaine Basin. Syst Appl Microbiol **34**:494–497.
2. **Hugenholtz P**, **Pitulle C**, **Hershberger KL**, **Pace NR**. 1998. Novel division level bacterial diversity in a Yellowstone hot spring. J Bacteriol **180**:366–376.
3. **Fu Q**, **Kuramochi Y**, **Fukushima N**, **Maeda H**, **Sato K**, **Kobayashi H**. 2015. Bioelectrochemical analyses of the development of a thermophilic biocathode catalyzing electromethanogenesis. Environ Sci Technol **49**:1225–1232.
4. **Nurk S**, **et al**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20**:714–737.
5. **Refojo PN**, **Ribeiro MA**, **Calisto F**, **Teixeira M**, **Pereira MM**. 2013. Structural composition of alternative complex III: Variations on the same theme. Biochim Biophys Acta **1827**:1378–1382.
6. **Borisov VB**, **Gennis RB**, **Hemp J**, **Verkhovsky MI**. 2011. The cytochrome bd respiratory oxygen reductases. Biochim Biophys Acta **1807**:1398–1413.
7. **Sutcliffe IC**. 2011. Cell envelope architecture in the Chloroflexi: a shifting frontline in a phylogenetic turf war. Environ Microbiol **13**:279–282.

Draft genome of *Ornatilinea apprima* 3PM-1, an anaerobic member of the Chloroflexi class *Anaerolineae*

Hemp J, Ward LM, Pace LA, Fischer WW. 2015. Draft genome sequence of Ornatilinea apprima P3M-1, an anaerobic member of the Chloroflexi class Anaerolineae. Genome Announc 3(6):e01353-15. DOI: 10.1128/genomeA.01353-15

Abstract

We report the draft genome of *Ornatilinea apprima* 3PM-1, a strictly anaerobic member of the Chloroflexi class *Anaerolineae*. This genome provides insight into the diversity of metabolism within the *Anaerolineae*, and the evolution of respiration within the Chloroflexi.

Genome Announcement

Ornatilinea apprima 3PM-1 was isolated in Siberia from a microbial mat in a wooden bathtub sourced with water from a 2775 m well (1). Closely related strains have been reported from fresh water lakes and rice paddy soils. *O. apprima* was physiologically characterized as a filamentous, non-motile, obligately anaerobic organotroph. It can ferment a wide range of polypeptides and carbohydrates, including microcrystalline cellulose. It grows optimally at 42-45 °C and pH 7.5-8.0 (1).

The genome of *Ornatilinea apprima* 3PM-1 (DSM 23815) was sequenced as part of a project to expand the phylogenetic breadth of Chloroflexi genomes. Genome sequencing was performed at Seqmatic using the Illumina MiSeq sequencing platform. SPAdes 3.1.1

(2) was used to assemble the genome. The genome was screened for contaminants based on sequence coverage, GC composition, and BLAST hits of conserved single copy genes. Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The draft genome is 4.41 Mb in size, assembled into 45 contigs. It encodes 3846 genes, 3347 CDSs, 2 16sRNA, 50 tRNAs, and 3 CRISPR arrays. It is estimated to be ~96% complete based on conserved single copy genes (107/111).

The majority of cultured Chloroflexi belong to the class *Chloroflexia*, which is composed of anoxygenic phototrophs and facultative aerobes (3). Even though > 70% of all Chloroflexi sequences present in 16S datasets belong to the *Anaerolineae* class (4), they are less well characterized. To date all described members are anaerobic fermentative organisms (5). Consistent with its description as an obligate anaerobe, *O. apprima* contains no genes for O_2 respiration. Even though it was classified as non-motile the genome does encode genes for flagella and chemotaxis, two traits not previously observed in *Anaerolineae*, suggesting that it is capable of motility. In addition no genes for LPS biosynthesis or outer-membrane proteins were found, which is consistent with the hypothesis that Chloroflexi have only one membrane (6).

References

1. **Podosokorskaya OA**, **Bonch-Osmolovskaya EA**, **Novikov AA**, **Kolganova TV**, **Kublanov IV**. 2013. Ornatilinea apprima gen. nov., sp. nov., a cellulolytic representative of the class Anaerolineae. Int J Syst Evol Microbiol **63**:86–92.

2. **Nurk S**, **Bankevich A**, **Antipov D**, **Gurevich AA**, **Korobeynikov A**, **Lapidus A**, **Prjibelski AD**, **Pyshkin A**, **Sirotkin A**, **Sirotkin Y**, **Stepanauskas R**, **Clingenpeel SR**, **Woyke T**, **McLean JS**, **Lasken R**, **Tesler G**, **Alekseyev MA**, **Pevzner PA**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20**:714–737.

3. **Overmann J**. 2008. Green Nonsulfur Bacteria. *In* Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd, Chichester, UK.

4. **Yamada T**, **Sekiguchi Y**. 2009. Cultivation of uncultured chloroflexi subphyla: significance and ecophysiology of formerly uncultured chloroflexi "subphylum i" with natural and biotechnological relevance. Microbes Environ **24**:205–216.

5. **Yamada T**, **Sekiguchi Y**, **Hanada S**, **Imachi H**, **Ohashi A**, **Harada H**, **Kamagata Y**. 2006. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. Int J Syst Evol Microbiol **56**:1331–1340.

6. **Sutcliffe IC**. 2011. Cell envelope architecture in the Chloroflexi: a shifting frontline in a phylogenetic turf war. Environ Microbiol **13**:279–282.

Draft genome of *Ardenticatena maritima* 110S, a thermophilic nitrate and iron-reducing member of the Chloroflexi class *Ardenticatenia*

Hemp J, Ward LM, Pace LA, Fischer WW. 2015. Draft genome sequence of Ardenticatena maritima 110S, a thermophilic nitrate- and iron-reducing member of the Chloroflexi class Ardenticatenia. Genome Announc 3(6):e01347-15. DOI: 10.1128/genomeA.01347-15

Abstract

We report the draft genome of *Ardenticatena maritima* 110S, the first sequenced member of the *Ardenticatenia* class of the Chloroflexi phylum. This thermophilic organism is capable of a range of physiologies, including aerobic respiration and iron reduction. It also encodes a complete denitrification pathway with a novel nitric oxide reductase.

Genome Announcement

Ardenticatena maritima 110S was originally isolated from an iron-rich coastal hydrothermal field in the Kirishima Volcanic Belt of Japan (1). Closely related strains have been reported from hot springs (2) and hydrothermal vents (3). *A. maritima* is a filamentous, non-motile organism that can facultatively reduce nitrate and iron (1). It grows optimally at 50-70 °C and pH 7.0 (range pH 5.5-8.0).

The genome of *Ardenticatena maritima* 110S (DSM 23922) was sequenced as part of a project to expand the phylogenetic breadth of Chloroflexi genomes. Genome sequencing was performed at Seqmatic using the Illumina MiSeq sequencing platform. SPAdes 3.1.1

(4) was used to assemble the genome. The genome was screened for contaminants based on sequence coverage, GC composition, and BLAST hits of conserved single copy genes. Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The draft genome is 3.62 Mb in size, assembled into 12 contigs. It encodes 3041 genes, 2578 CDSs, 2 16sRNA, 46 tRNAs, and 10 CRISPR arrays. It is estimated to be ~99% complete based on conserved single copy genes (111/111).

Analysis of the *A. maritima* genome revealed the presence of many genes responsible for its physiological breadth. *A. maritima* encodes for a branched aerobic respiratory chain, including: Complex I (NADH dehydrogenase), Complex II (succinate dehydrogenase), Complex III (cytochrome *bc* complex) and three oxygen reductases (A and B-family heme-copper oxygen reductases and *bd* oxidase). B-family heme-copper oxygen reductases are commonly found in aerobic thermophiles, enabling growth with the low oxygen levels found in thermal systems (5). *A. maritima* also encodes a complete dentrification pathway composed of nitrate reductase (NapA), nitrite reductase (NirK), a novel nitric oxide reductase (eNOR) (6), and nitrous oxide reductase (NosZ). Interestingly the eNOR is found in an operon with NirK, suggesting that these genes are co-regulated. No genes were found for either LPS biosynthesis or outermembrane proteins, consistent with the proposal that Chloroflexi only have one membrane (7).

References

1. **Kawaichi S**, **Ito N**, **Kamikawa R**, **Sugawara T**, **Yoshida T**, **Sako Y**. 2013.

Ardenticatena maritima gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum “Chloroflexi” isolated from an iron-rich coastal hydrothermal field, and description of Ardenticatenia classis nov. Int J Syst Evol Microbiol **63**:2992–3002.

2. **Kormas KA**, **Tamaki H**, **Hanada S**. 2009. Apparent richness and community composition of Bacteria and Archaea in geothermal springs. Aquatic Microbial ….

3. **Meyer-Dombard DR**, **Amend JP**, **Osburn MR**. 2013. Microbial diversity and potential for arsenic and iron biogeochemical cycling at an arsenic rich, shallow-sea hydrothermal vent (Tutum Bay, Papua New Guinea). Chemical Geology **348**:37–47.

4. **Nurk S**, **Bankevich A**, **Antipov D**, **Gurevich AA**, **Korobeynikov A**, **Lapidus A**, **Prjibelski AD**, **Pyshkin A**, **Sirotkin A**, **Sirotkin Y**, **Stepanauskas R**, **Clingenpeel SR**, **Woyke T**, **McLean JS**, **Lasken R**, **Tesler G**, **Alekseyev MA**, **Pevzner PA**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20**:714–737.

5. **Chang H-Y**, **Hemp J**, **Chen Y**, **Fee JA**, **Gennis RB**. 2009. The cytochrome ba3 oxygen reductase from Thermus thermophilus uses a single input channel for proton delivery to the active site and for proton pumping. Proc Natl Acad Sci USA **106**:16169–16173.

6. **Hemp J**, **Gennis RB**. 2008. Diversity of the heme-copper superfamily in archaea: insights from genomics and structural modeling. Results Probl Cell Differ **45**:1–31.

7. **Sutcliffe IC**. 2011. Cell envelope architecture in the Chloroflexi: a shifting frontline in a phylogenetic turf war. Environ Microbiol **13**:279–282.

A p p e n d i x 5

CROWN GROUP OXYPHOTOBACTERIA POSTDATE THE RISE OF OXYGEN

Patrick M. Shih, James Hemp, Lewis M. Ward, Nicholas J. Matzke, and Woodward W. Fischer. "Crown group Oxyphotobacteria postdate the rise of oxygen." Geobiology 15.1 (2017): 19-29. DOI: 10.1111/gbi.12200

Abstract:

The rise of oxygen *ca.* 2.3 billion years ago (Ga) is the most distinct environmental transition in Earth history. This event was enabled by the evolution of oxygenic photosynthesis in the ancestors of Cyanobacteria. However, longstanding questions concern the evolutionary timing of this metabolism, with conflicting answers spanning more than one billion years. Recently, knowledge of the Cyanobacteria phylum has expanded with the discovery of non-photosynthetic members, including a closely related sister group termed Melainabacteria, with the known oxygenic phototrophs restricted to a clade recently designated Oxyphotobacteria. By integrating genomic data from the Melainabacteria, cross-calibrated Bayesian relaxed molecular clock analyses show that crown group Oxyphotobacteria evolved *ca.* 2.0 billion years ago (Ga), well after the rise of atmospheric dioxygen. We further estimate the divergence between Oxyphotobacteria and Melainabacteria *ca.* 2.5-2.6 Ga, which—if oxygenic photosynthesis is an evolutionary synapomorphy of the Oxyphotobacteria—marks an upper limit for the origin of oxygenic

photosynthesis. Together these results are consistent with the hypothesis that oxygenic photosynthesis evolved relatively close in time to the rise of oxygen.

Introduction:

Oxygenic photosynthesis was responsible for the most profound environmental shift in Earth history: the rise of oxygen. It was long recognized that this metabolism evolved in the Cyanobacteria phylum, and that this unique ability was a necessary precondition for the rise of oxygen at *ca*. 2.35 Ga [1]. However the evolutionary origins of Cyanobacteria remains uncertain, due to many conflicting lines of evidence ranging from microfossils [2], geochemical data [3-6], biomarkers [7], and geochemical models [8, 9]. These different proxies for Cyanobacteria provide divergence estimates that span more than one billion years of Earth history. This quandary leaves a major gap in knowledge regarding the O_2 cycle; consequently it remains unclear whether the rise of oxygen was directly related to the evolution of oxygenic photosynthesis [4, 10], or alternatively driven by a change in Earth's geophysical processes [8, 9].

The origins of Cyanobacteria has been intensely debated over the last half-century. Historically, hypotheses have been grounded in observations from the fossil and sedimentary rock records. However, the validity of some of the landmark studies, which have become core to many fundamental assumptions of the field, has been called into question, specifically concerning the biogenicity and taxonomic affinity of microfossils [11-13] and the recent reevaluation of Archean molecular fossils as sample contaminants

[14]. Because of the paucity of unequivocal evidence from early Precambrian sedimentary rocks, phylogenetics and comparative genomics provide an independent approach to constrain the evolutionary timing of Cyanobacteria, offering a useful point of comparison with the geological and fossil records [15].

A natural way to reconstruct and evaluate the evolution of oxygenic photosynthesis by Cyanobacteria from comparative biology is to identify their closest living relatives, which until recently remained a mystery. The notion that all Cyanobacteria evolved from a common photosynthetic ancestor hinged upon limited sampling of their extant diversity. Adding to the uncertainty, previous phylogenetic efforts to distinguish bacterial phylum-level relationships and identify the closest relative to Cyanobacteria have yielded conflicting results [16-19]. However, rapidly growing genomic and metagenomic datasets have substantially added to our understanding of the Cyanobacteria phylum. 16S rDNA surveys first described a substantial diversity of microbes related to oxygenic Cyanobacteria from aphotic environments [20]. Recently, genomes from a wide range of aphotic environments, such as gut microbiomes and groundwater samples, were analyzed that revealed the presence of a close sister group to oxygenic Cyanobacteria, named Melainabacteria, filling in an important gap in the diversity and evolution of the Cyanobacteria phylum [20-24]. The close evolutionary relationships of Melainabacteria and previously known Cyanobacteria fit with widely applied guidelines using 16S rDNA sequence identity for inclusion in microbial phyla [22]; this is a similar degree of sequence identity, for example, as that observed between *Escherichia coli* and *Pseudomonas*

aeruginosa—two common and well studied members of the class γ-Proteobacteria in the phylum Proteobacteria [25]. The placement of Melainabacteria as a close sister group to all known phototrophic Cyanobacteria is objective and reproducible on the basis a myriad of different gene and protein comparisons [21-23, 26, 27]. Additionally, the known oxygenic Cyanobacteria are much more closely related to the Melainabacteria than they are to any of the currently known members of other phyla capable of anoxygenic phototrophy, which are placed at far greater evolutionary distances, with numerous non-phototrophic groups interspersed between them [26]. These data illustrated that the Cyanobacteria phylum hosts a greater degree of physiological diversity than previously recognized – a condition similar to all the other known phototrophic phyla, which contain both phototrophic and non-phototrophic members [26].

With the discovery of a substantial diversity of close-living relatives to the oxygenic Cyanobacteria, it has been proposed to update the systematics of the Cyanobacteria phylum by adding a Melainabacteria class and relegating the oxygenic Cyanobacteria to the Oxyphotobacteria class [22]. For ease of discussion, we refer to these two sister clades using the following nomenclature: Oxyphotobacteria (class *Oxyphotobacteria* containing all known oxygenic Cyanobacteria) and Melainabacteria, wherein the Melainabacteria class to date includes four orders *Gastranaerophilales* [20, 21], *Obscuribacterales*, *Vampirovibrionales*, *Caenarcaniphilales* [22] (Figure 1). Interestingly, out of all the Melainabacteria genomes sequenced, none contain any genes associated with photosynthesis, and many lack genes necessary for aerobic or anaerobic respiration [21,

22]. Despite substantial environmental and genomic diversity, due to the current lack of observed photosynthetic basal lineages it is important to consider the hypothesis that oxygenic photosynthesis is a derived, and perhaps relatively recent, feature of the Cyanobacteria phylum (Figure 1). Thus, the newly identified Melainabacteria can add useful phylogenetic information in the sparsely covered regions closer to the base of the Cyanobacterial phylum.

To estimate the origin of Oxyphotobacteria, molecular clock studies have typically used either molecular fossils [28, 29] or cyanobacterial-like microfossils [30-32] as calibration constraints. It was once common to interpret microfossils as specific living lineages on the basis of morphological traits; however, phylogenetic analyses have revealed multiple independent acquisitions and widespread convergences of many of the classical cyanobacterial morphotypes, such as baeocystous and filamentous cells [33, 34], highlighting classic challenges in assigning microfossils to extant clades [35]. Moreover, 2-methylhopane molecular fossils, once thought to have been specific to the oxygenic Cyanobacteria, appear to have evolved in other phyla [36] and may not offer unique calibration constraints, concerns about syngeneity aside [14]. A number of studies have used the rise of oxygen as a calibration point for the minimal age of Oxyphotobacteria [28, 30, 31, 37]—this placement implicitly assumes that crown group Oxyphotobacteria were responsible for the rise of oxygen. However, it is equally plausible that extinct lineages existing before the most recent common ancestor of Oxyphotobacteria (*i.e.*, stem lineages) sourced the O_2 fluxes connected with the rise of oxygen. In order to test these assumptions,

it is useful to relax these constraints. Instead, valuable evolutionary insights into the origins of Oxyphotobacteria comes from endosymbiosis, which lends constraints from the fossil records from plants and algae – characteristics that most other bacterial phyla do not share. Additionally, by incorporating new molecular data from Melainabacteria taxa molecular clock analyses can now obtain a tighter estimate of the divergence time of Oxyphotobacteria.

Here, we revisit this evolutionary problem by performing cross-calibrated Bayesian relaxed molecular clock analyses with increased taxonomic sampling encompassing Melainabacteria, Oxyphotobacteria, plastids, and mitochondria. Distinct from previous cross-calibration efforts, we expand upon the technique by using a concatenated dataset composed of slowly evolving proteins and genes found in both plastids and mitochondria, which permit calibrations to be used multiple times across a phylogenetic tree [38].

Methods:

Generation of Concatenated Dataset. Sequences from subunits of ATP synthase, the ribosomal large subunit, the ribosomal small subunit, elongation factor Tu, and 16S rDNA were gathered. The protein sequences gathered from ATP synthase machinery consisted of AtpA, AtpB, AtpE, AtpF, AtpH, and AtpI. The ribosomal protein subunits collected for this study were Rpl2, Rpl16, Rps3, and Rps12. All sequences and their corresponding accessions are listed in Table S1. Sequences were collected to maximize the coverage of the Oxyphotobacteria and plastid-bearing eukaryotes. Plant mitochondrial genomes were

used along with α-Proteobacteria to serve as both an appropriate outgroup and enable cross-calibration between the corresponding plant mitochondrial and plastid lineages. Alignments for each protein or nucleotide family were performed using the –maxiterate strategy in the alignment program MAFFT [39], and then concatenated to generate the final dataset. The dataset was partitioned into two parts: concatenated protein and 16S nucleotide sequences, respectively.

Age Calibrations. Dating calibration priors were primarily chosen to avoid biasing analyses with the introduction of controversial microfossil occurrences, instead relying on well-accepted divergence times of plant fossils estimated. The posterior of divergence times from the comprehensive molecular clock analysis of land plants by Smith et al. were used as priors in this study [40]. A summary of all constraints used is described in Table 1. Normal distributions of 217 ± 40, 327 ± 30, 432 ± 30, and 477 ± 70 Ma were used as divergence time calibration points for Angiospermae, Gymnospermatophyta, Tracheophyta, and land plants, respectively. Importantly, the use of land plant divergence events enabled the cross-calibration between divergence events happening simultaneously in lineages that contain both plastids and mitochondria. In BEAST runs incorporating the fossil *Bangiomorpha pubescens* were constrained using a normal distribution between 1174-1222 Ma, as done in Yoon et al [41]. Strict geochronologic constraints for the strata from which these fossils were found – the Hunting Formation on Somerset Island, Nunavut, Canada – are between *ca*. 1200 and *ca*. 900 Ma. We examined a constraint associated with the oldest estimate placed between 1174 and 1222 Ma. This was chosen in

part to provide a conservative (*i.e.*, oldest possible) estimate on the divergence of crown group Oxyphotobacteria [42, 43]. Because the timing of the evolution of oxygenic photosynthesis remains controversial based on geological observations, we used a uniform prior spanning between 2400 and 3000 Ma and tested this prior on the two different nodes that represent 1) the divergence between Melainabacteria and Oxyphotobacteria and 2) the radiation of crown group Oxyphotobacteria. The "Rise of Oxygen" constraint provides a minimum age of 2400 Ma [44], marking the oldest ages hypothesized for the rise of oxygen, and an upper age of 3000 Ma as suggested by various geological studies [5, 45]. A uniform distribution enables the Markov chain Monte Carlo (MCMC) search to agnostically explore with no initial bias before ultimately converging onto a date within the provided hard upper and lower bounds of 2400-3000 Ga. Finally, a uniform prior between 3800 and 2400 Ma was used as a calibration for the last common ancestor of all taxa used in this study. The large range was used to permit an unbiased and largely unrestricted constraint, which assumes that oxygenic photosynthesis must predate the rise of oxygen and that these taxa had not diverged prior to the Late Heavy Bombardment *ca*. 3800 Ma [46].

Molecular Clock Analysis. Dated phylogenies were estimated using the program BEAST [47] using the CIPRES Science Gateway v. 3.3 server [48]. Cross-calibrated analyses were coded into BEAST XML files as previously described [38]. For each macrofossil that provides an age calibration, the same date is applied to every node in the molecular phylogeny that corresponds to the speciation event. Thus, if paralogs of a gene are found in

the nucleus, chloroplast, and mitochondrion, one calibration fossil supplies calibration distributions for several nodes in the gene phylogeny. The CpREV model was chosen as the best-fitting amino acid substitution model based on ProtTest analysis [49]. The CpREV model was used for the concatenated amino acid sequences and the 16S rDNA sequences used the GTR+G model in accordance to the 16S rDNA molecular clock study using BEAST by Schirrmeister et al [30]. A variety of BEAST runs were calculated using different combinations of date calibration priors, described below. For all runs, we ran five MCMC chains for the maximum limit allowed on the CIPRES Science Gateway, sampling every 10,000th generation. On average, more than 20 million generations were collected from each MCMC chain, and the initial 5 million generations were discarded as burnin (for exact numbers of post-burnin generations reported in Table 2). Maximum clade credibility (MCC) trees were generated using TreeAnnotator v1.7.5. All BEAST runs are summarized in Table 2. In general, the resulting chronograms of runs T64, T65, T68, and T69 were consistent with the overall phylogenomic analyses revealing the sister relationship between Oxyphotobacteria and Melainabacteria described by both Soo et al [22] and Di Rienzi et al [21]. In contrast, the topology of T72 and T73 placed Melainabacteria sister to α-Proteobacteria.

Regression Analysis of Node Age Uncertainty. The 95% Highest Posteriori Density (HPD) widths of node dates from the MCC trees generated by each analysis were extracted and plotted in R. For each pair of analyses, the list of HPD widths was reduced to the set of nodes common between the trees (the trees had mostly common nodes), and

the HPD widths were regressed on each other and plotted. A 1:1 line was also plotted for visual comparison. To statistically test for deviations from a 1:1 line (indicating that one analysis had a higher or lower amount of uncertainty in node dates than the other), the expected 1:1 relationship was subtracted and the regression was repeated; the resulting p-value on slope (printed on each plot) thus tests for significant deviation from the expected 1:1 relationship.

Expanded Phylogeny of Melainabacteria and Oxyphotobacteria. 16S rDNA sequences from members of the Cyanobacteria phylum (including the Oxyphotobacteria, Melainabacteria, and ML635J-21 classes) greater than 1250 base pairs in length were acquired from the Silva database [50]. Alignments were performed using the Infernal aligner [51] implemented by the Ribosomal Database Project [52]. Multiple phyla (Chloroflexi, OP11, Armatimonodetes, BD1-5, OD1, SHA-109, SR1, TM7, WWE3, and WS6) were used as out groups for the 16s rDNA phylogeny. Phylogenetic analyses were performed at the CIPRES Science Gateway using RAxML under the GTR model and default parameters for nucleotides. Trees were imaged using Interactive Tree of Life software [53].

Phylogenetic Analysis of O_2 reductases. Sequences for the A and C-family heme-copper O_2 reductases were acquired from public genomic and metagenomic databases and aligned using MAFFT. The resulting alignments were manually curated using structural and biochemical data. Phylogenetic analyses were performed at the CIPRES Science Gateway

using RAxML [54] under the LG model. The A and C-family O_2 reductases were each mid-point rooted due to uncertainty in the true position of the roots for these related protein families. Trees were imaged using Interactive Tree of Life software [53].

Results and Discussion:

Crown group Oxyphotobacteria postdate the rise of oxygen.

We have previously shown that cross-calibrated of Bayesian relaxed molecular clock analyses improve estimates of timing of evolutionary events, and can be useful to date Precambrian divergences [38]. Here, we built upon this technique by constructing and analyzing a dataset of slowly evolving proteins and genes found in both plastids and mitochondria, which permit fossil calibrations to be used multiple times across a phylogenetic tree. This included proteins from ATP synthase, the ribosomal large subunit, the ribosomal small subunit, and elongation factor Tu, as well as 16S rDNA sequences from Oxyphotobacteria, Melainabacteria, plastids, α-Proteobacteria, and plant mitochondria (Table S1-S3).

By expanding the cross-calibrated analysis to include Melainabacteria, we estimate reveal the divergence of crown group Oxyphotobacteria at *ca*. 2.0 Ga, postdating the rise of oxygen by at least 300 million years (Table 3, Table S2). A useful way to test the robustness of this estimate is by examining the results under different combinations of age calibration constraints. Permutations of the two deepest Proterozoic calibrations—the "Rise of Oxygen" constraint and the *Bangiomorpha* fossil constraint—all show a

Paleoproterozoic radiation of crown group Oxyphotobacteria, ranging in time between 2.071-1.741 Ga (Table 3, Figure 2, Figure S1-S4). This result contrasts with previous molecular clock studies that explicitly constrained the most recent common ancestor of all Oxyphotobacteria to the rise of oxygen, under the assumption that the radiation of the crown group was responsible for the rise of oxygen [28, 30, 31, 37]. However, this finding is in agreement with the fossil record, where the first widely accepted fossil Cyanobacteria (*Eoentophysalis sp*.) occur in shallow marine carbonate strata at *ca*. 1.9 Ga [55]. In contrast, when the "Rise of Oxygen" constraint was placed on the radiation of crown Oxyphotobacteria, the Melainabacteria clade instead became sister to the α-Proteobacteria/mitochondria clade – a result in conflict with all phylogenetic studies to date [20-22], which suggested that this calibration set poorly fit the data (Figure S5 & S6). Moreover, comparisons of node age uncertainty between BEAST analyses supported placement of the "Rise of Oxygen" constraint on the divergence of Melainabacteria and Oxyphotobacteria over the radiation of crown Oxyphototbacteria (Figure 3).

Dating the divergence between Oxyphotobacteria and Melainabacteria.

Next, we examined the hypothesis that Oxyphotobacteria and Melainabacteria diverged prior to the rise of oxygen. With a sister group relationship from current genomic data there is a degree of ambiguity regarding metabolic reconstruction of the last common ancestor of these clades. Two reasonable evolutionary scenarios exist (Figure 4). 1) The common ancestor to both Oxyphotobacteria and Melainabacteria was aphototrophic and oxygenic photosynthesis evolved after their divergence in the Oxyphotobacteria lineage. If correct,

the divergence of these clades would mark an upper age limit on the evolution of oxygenic photosynthesis. 2) Oxygenic photosynthesis evolved prior to the divergence of Oxyphotobacteria and Melainabacteria and was subsequently lost from the Melainabacteria clade. The BEAST results can, in principle, help select between these possibilities. For example, if we observe that they diverged after the rise of oxygen, that result would support a loss of photosynthesis from the Melainabacteria (scenario 2).

Our cross-calibrated analyses estimate the divergence of Oxyphotobacteria and Melainabacteria to have occurred *ca*. 2.5-2.6 Ga (Table 3). This result is consistent with the hypothesis that stem group Oxyphotobacteria evolved oxygenic photosynthesis after their divergence from the Melainabacteria, relatively close in time to the rise of oxygen [10, 56, 57]. This is supported by the observation that the Oxyphotobacteria are nested within a diverse clade of organisms from aphotic environments (Figure 1). Furthermore, no characterized Melainabacteria genomes encode genes involved in photosynthesis, whereas oxygenic photosynthesis is conserved within all extant Oxyphotobacteria except for a few clear cases of losses in obligate symbionts [58, 59]. Additional support for scenario 1 is provided by the contrasting evolutionary histories of other aspects of high potential metabolism (e.g. aerobic respiration) between Oxyphotobacteria and Melainabacteria. As mentioned above, no Melainabacteria contain genes for photosynthesis, but all extant Oxyphotobacteria and some Melainabacteria are able to perform oxidative phosphorylation [22], leading to the question of when this ability was acquired within the Cyanobacteria phylum. The F_0F_1-ATP synthase is the only phylogenetically congruent protein shared

between the Oxyphotobacteria and Melainabacteria used for oxidative phosphorylation. Phylogenetic analyses of the other respiratory proteins described below illustrate that they were acquired after the divergence of the two clades, suggesting that aerobic respiration and high potential metabolism was acquired independently in the Oxyphotobacteria and Melainabacteria.

For aerobic respiration, Oxyphotobacteria use evolutionarily conserved A-family O_2 reductases; these proteins display a characteristically low affinity for O_2, and imply that Oxyphotobacteria acquired their A-family O_2 reductases after the evolution of oxygenic photosynthesis and oxygenation of their environment. In contrast, some Melainabacteria in the *Obscuribacterales*, *Vampirovibrionales*, and *Caenarcaniphilales* orders are capable of aerobic respiration using C-family O_2 reductases. Notably, the C-family O_2 reductases are only very distantly related to the A-family O_2 reductases from Oxyphotobacteria [60], and occur within an operon with genes for Complex III. This suggests that the Melainabacteria capable of aerobic respiration acquired this ability by lateral gene transfer after the evolution of oxygenic phototrophy (Figure S7 & S8). Thus aerobic respiration was acquired independently at least twice within this phylum (once each in the Oxyphotobacteria and Melainabacteria). Oxyphotobacteria and Melainabacteria also utilize two substantially different Complex IIIs for aerobic respiration. The Oxyphotobacteria use a b_6f complex that functions for both phototrophy and aerobic respiration. The b_6f complex has a split cytochrome *b*, an additional cytochrome c_i near the Q_i plastoquinol binding site, and a novel cytochrome *f*. The aerobic Melainabacteria instead have a Complex III with a

full-length cytochrome *b* and no cytochrome *c*-containing subunit [22]. Thus, current genomic data supports a scenario in which aerobic respiration evolved after the origin of oxygenic photosynthesis in this phylum. Again this is more consistent with scenario 1: Oxygenic photosynthesis evolved after the divergence of Oxyphotobacteria from Melainabacteria, with the independent acquisition of aerobic respiration within these groups after the rise of oxygen. Although several lines of evidence favor scenario 1 over scenario 2, a formal test of the hypothesis described in scenario 2 will require genomic data from yet more basal members of this bacterial phylum (e.g. members of ML635J-21).

Compared with geological data, a Neoarchean 2.5-2.6 Ga divergence result for Oxyphotobacteria from their closest living relatives is also consistent with observations of a transitional photosystem using Mn prior to water-splitting at 2.415 Ga [61], as well as a number of studies suggesting small photosynthetic fluxes of O_2 in Neoarchean environments prior to the rise of oxygen [62, 63]. On the other hand, this result is discordant with hypotheses that stretch Oxyphotobacteria deep into Mesoarchean and older intervals [3, 5, 45], and raises the possibility that if those geochemical data are interpreted correctly, they may reflect non-cyanobacterial sources of O_2 or other oxidants [26, 64, 65].

Evolution of photosynthesis in Oxyphotobacteria.

Oxygenic photosynthesis would eventually become the core engine of the carbon cycle, but this need not have occurred synchronously with the evolution of water-splitting. All analyses show several hundred million years between the divergence of Melainabacteria

and Oxyphotobacteria, and the radiation of crown group Oxyphotobacteria (Figure 2). This length of time is important because a large number of evolutionary characters (including a number of multi-subunit complexes) are associated with extant oxygenic photosynthetic organisms (and missing from Melainabacteria), including photosystems I and II, the b_6f complex, RuBisCO and the Calvin cycle, and aerobic respiration – all are synapomorphies of Oxyphotobacteria [66]. It is unlikely that these traits all evolved at the same time, and comparisons of genomic data from Oxyphotobacteria, Melainabacteria, and additional basal clades can help ordinate the relative timing of the appearance of these characters.

The new dates for the divergence of Oxyphotobacteria from Melainabacteria, and the crown group radiation of Oxyphotobacteria, also provide a test of certain models for the evolution of phototrophy. As mentioned above, although all phototrophy is built around the same common molecular machinery (reaction centers, chlorins, and complex III or alternative complex III), it is also clear that the six phyla that are known to contain phototrophic members today are not closely related to one another; and consequently, lateral gene transfer was an important evolutionary vector for the sparse and scattered distribution of phototrophy observed today [26, 67]. In one class of hypotheses for origin of phototrophy—termed the fusion hypothesis—the Type I and Type II reaction centers evolved in different lineages (none of which must remain extant), and that ancestral Oxyphotobacteria then ultimately acquired both types of reaction centers by lateral gene transfer [26, 68]. If it is correct that oxygenic photosynthesis is derived within the Cyanobacteria phylum, then our estimates of 2.5-2.6 Ga for the divergence between

Oxyphotobacteria and Melainabacteria support the fusion hypothesis, because geological observations highlight that anoxygenic phototrophy was present at *ca*. 3.4 Ga [69].

Conclusions:

We presented the first molecular clock estimates for key divergences in the Cyanobacteria phylum that include data from close-living non-phototrophic relatives, but do not employ Archean lipid biomarkers, putative Cyanobacteria microfossils, or the rise of oxygen as calibration constraints. The results from cross-calibrated molecular analyses that include genomic data from the newly discovered Melainabacteria suggest that all known oxygenic Cyanobacteria did not appear until relatively late in Earth history [10, 26]. Estimates of the timing of divergence between Oxyphotobacteria and Melainabacteria are ~ 2.5-2.6 Ga. If it can be uniquely determined that oxygenic photosynthesis is a synapomorphy of the Oxyphotobacteria – and not simply loss or losses of both phototrophy and respiration from the Melainabacteria – this constraint would correspond to a maximum age of oxygenic photosynthesis (Figure 4, scenario 1). Current data best support this interpretation, but a formal test of this hypothesis awaits more genomic data from basal members of this phylum.

The results also consistently highlight that crown group oxygenic Cyanobacteria substantially postdate the rise of oxygen—which might instead be attributed to O_2 fluxes provided by organisms belonging to stem lineages. The results also consistently highlight that crown group Oxyphotobacteria (and thus all currently known Cyanobacteria capable of

oxygenic photosynthesis) substantially postdate the rise of oxygen—which might instead be attributed to O_2 fluxes provided by organisms belonging to stem lineages. Many previous studies have placed this crown group divergence on or prior to the rise of oxygen, but based on our results that show this assumption is not well supported; we suggest that it may be useful for future molecular clock studies to examine analyses with and without it. Further comparative genomic analyses reveal that aerobic respiration was acquired at least twice in the Cyanobacteria phylum after the evolution of oxygenic phototrophy.

Ultimately, our understanding of the major metabolic innovations that underpin Earth's biogeochemical cycles is limited to interpretations of the geochemical and fossil data in the context of frameworks built by knowledge of extant biology. This study highlights the opportunity to leverage the exponentially growing amount of genomic and metagenomic data in efforts to update and improve the quality of these frameworks.

References:

1. Luo, G., Ono, S., Beukes, N.J., Wang, D.T., Xie, S., and Summons, R.E. (2016). Rapid oxygenation of Earth's atmosphere 2.33 billion years ago. Science Advances *2*.
2. Schopf, J.W., and Packer, B.M. (1987). Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia. Science *237*, 70-73.

3. Rosing, M.T., and Frei, R. (2004). U-rich Archaean seafloor sediments from Greenland—indications of >3700 Ma oxygenic photosynthesis. Earth Planet. Sci. Lett. *217*, 237-244.

4. Kopp, R.E., Kirschvink, J.L., Hilburn, I.A., and Nash, C.Z. (2005). The Paleoproterozoic snowball Earth: A climate disaster triggered by the evolution of oxygenic photosynthesis. Proc Natl Acad Sci *102*, 11131-11136.

5. Crowe, S.A., Dossing, L.N., Beukes, N.J., Bau, M., Kruger, S.J., Frei, R., and Canfield, D.E. (2013). Atmospheric oxygenation three billion years ago. Nature *501*, 535-538.

6. Johnson, J.E., Gerpheide, A., Lamb, M.P., and Fischer, W.W. (2014). O2 constraints from Paleoproterozoic detrital pyrite and uraninite. Geol Soc Am Bull.

7. Brocks, J.J., Logan, G.A., Buick, R., and Summons, R.E. (1999). Archean Molecular Fossils and the Early Rise of Eukaryotes. Science *285*, 1033-1036.

8. Holland, H.D. (2009). Why the atmosphere became oxygenated: A proposal. Geochim Cosmochim Acta *73*, 5241-5255.

9. Kump, L.R., and Barley, M.E. (2007). Increased subaerial volcanism and the rise of atmospheric oxygen 2.5 billion years ago. Nature *448*, 1033-1036.

10. Ward, L., Kirschvink, J., and Fischer, W. (2015). Timescales of Oxygenation Following the Evolution of Oxygenic Photosynthesis. Orig Life Evol Biosph, 1-15.

11. Brasier, M.D., Green, O.R., Jephcoat, A.P., Kleppe, A.K., Van Kranendonk, M.J., Lindsay, J.F., Steele, A., and Grassineau, N.V. (2002). Questioning the evidence for Earth's oldest fossils. Nature *416*, 76-81.

12. Butterfield, N.J. (2015). Proterozoic photosynthesis – a critical review. Palaeontology *58*, 953-972.

13. Knoll, A.H. (2003). The geological consequences of evolution. Geobiology *1*, 3-14.

14. French, K.L., Hallmann, C., Hope, J.M., Schoon, P.L., Zumberge, J.A., Hoshino, Y., Peters, C.A., George, S.C., Love, G.D., Brocks, J.J., et al. (2015). Reappraisal of hydrocarbon biomarkers in Archean rocks. Proc Natl Acad Sci *112*, 5915-5920.

15. Shih, P.M. (2015). Photosynthesis and early Earth. Curr Biol *25*, R855-R859.

16. Zhaxybayeva, O., and Gogarten, J.P. (2002). Bootstrap, Bayesian probability and maximum likelihood mapping: exploring new tools for comparative genome analyses. BMC Genomics *3*, 4.

17. Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N.N., Kunin, V., Goodwin, L., Wu, M., Tindall, B.J., et al. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. Nature *462*, 1056-1060.

18. Wolf, Y., Rogozin, I., Grishin, N., Tatusov, R., and Koonin, E. (2001). Genome trees constructed using five different approaches suggest new major bacterial clades. BMC Evol Biol *1*, 8.

19. Daubin, V., Gouy, M., and Perrière, G. (2002). A Phylogenomic Approach to Bacterial Phylogeny: Evidence of a Core of Genes Sharing a Common History. Genome Res *12*, 1080-1090.

20. Ley, R.E., Bäckhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D., and Gordon, J.I. (2005). Obesity alters gut microbial ecology. Proc Natl Acad Sci *102*, 11070-11075.

21. Di Rienzi, S.C., Sharon, I., Wrighton, K.C., Koren, O., Hug, L.A., Thomas, B.C., Goodrich, J.K., Bell, J.T., Spector, T.D., Banfield, J.F., et al. (2013). The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. eLife *2:e01102*.

22. Soo, R.M., Skennerton, C.T., Sekiguchi, Y., Imelfort, M., Paech, S.J., Dennis, P.G., Steen, J.A., Parks, D.H., Tyson, G.W., and Hugenholtz, P. (2014). An Expanded Genomic Representation of the Phylum Cyanobacteria. Genome Biol Evol *6*, 1031-1045.

23. Johnson, J.E., Webb, S.M., Thomas, K., Ono, S., Kirschvink, J.L., and Fischer, W.W. (2013). Reply to Jones and Crowe: Correcting mistaken views of sedimentary geology, Mn-oxidation rates, and molecular clocks. Proc Natl Acad Sci *110*, E4119-E4120.

24. van der Lelie, D., Taghavi, S., McCorkle, S.M., Li, L.-L., Malfatti, S.A., Monteleone, D., Donohoe, B.S., Ding, S.-Y., Adney, W.S., Himmel, M.E., et al. (2012). The Metagenome of an Anaerobic Microbial Community Decomposing Poplar Wood Chips. PLoS ONE *7*, e36740.

25. Hugenholtz, P., Goebel, B.M., and Pace, N.R. (1998). Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. J Bacteriol *180*, 4765-4774.

26. Fischer WW, H.J., Johnson JE (2016). Evolution of oxygenic photosynthesis. Annu Rev Earth Planet Sci *44*.

27. Fischer WW, H.J., Valentine JS (2016). How did life survive Earth's Great Oxidation? Curr Opin Chem Biol *in review*.

28. Battistuzzi, F., Feijao, A., and Hedges, S.B. (2004). A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. BMC Evol Biol *4*, 44.

29. Battistuzzi, F.U., and Hedges, S.B. (2009). A Major Clade of Prokaryotes with Ancient Adaptations to Life on Land. Mol Biol Evol *26*, 335-343.

30. Schirrmeister, B.E., de Vos, J.M., Antonelli, A., and Bagheri, H.C. (2013). Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event. Proc Natl Acad Sci.

31. Sánchez-Baracaldo, P., Ridgwell, A., and Raven, John A. (2014). A Neoproterozoic Transition in the Marine Nitrogen Cycle. Curr Biol *24*, 652-657.

32. Tomitani, A., Knoll, A.H., Cavanaugh, C.M., and Ohno, T. (2006). The evolutionary diversification of cyanobacteria: Molecular–phylogenetic and paleontological perspectives. Proc Natl Acad Sci *103*, 5442-5447.

33. Shih, P.M., Wu, D., Latifi, A., Axen, S.D., Fewer, D.P., Talla, E., Calteau, A., Cai, F., Tandeau de Marsac, N., Rippka, R., et al. (2013). Improving the coverage

of the cyanobacterial phylum using diversity-driven genome sequencing. Proc Natl Acad Sci *110*, 1053-1058.

34. Turner, S., Pryer, K.M., Miao, V.P.W., and Palmer, J.D. (1999). Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. J Eukaryot Microbiol *46*, 327-338.

35. Knoll, A.H., and Golubic, S. (1992). Proterozoic and Living Cyanobacteria. In Early Organic Evolution, M. Schidlowski, S. Golubic, M. Kimberley, D. McKirdy, Sr. and P. Trudinger, eds. (Springer Berlin Heidelberg), pp. 450-462.

36. Ricci, J.N., Michel, A.J., and Newman, D.K. (2015). Phylogenetic analysis of HpnP reveals the origin of 2-methylhopanoid production in Alphaproteobacteria. Geobiology, DOI: 10.1111/gbi.12129.

37. Falcon, L.I., Magallon, S., and Castillo, A. (2010). Dating the cyanobacterial ancestor of the chloroplast. ISME J *4*, 777-783.

38. Shih, P.M., and Matzke, N.J. (2013). Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. Proc Natl Acad Sci *110*, 12355-12360.

39. Katoh, K., Kuma, K.-i., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res *33*, 511-518.

40. Smith, S.A., Beaulieu, J.M., and Donoghue, M.J. (2010). An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. Proc Natl Acad Sci.

41. Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol *21*, 809-818.

42. Butterfield, N.J., Knoll, A.H., and Swett, K. (1988). Exceptional preservation of fossils in an Upper Proterozoic shale. Nature *334*, 424-427.

43. Butterfield, N.J., Knoll, A.H., and Swett, K. (1990). A bangiophyte red alga from the Proterozoic of arctic Canada. Science *250*, 104-107.

44. Hoffman, P.F. (2013). The Great Oxidation and a Siderian snowball Earth: MIF-S based correlation of Paleoproterozoic glacial epochs. Chem Geol *362*, 143-156.

45. Planavsky, N.J., Asael, D., Hofmann, A., Reinhard, C.T., Lalonde, S.V., Knudsen, A., Wang, X., Ossa Ossa, F., Pecoits, E., Smith, A.J.B., et al. (2014). Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. Nature Geosci *7*, 283-286.

46. Cohen, B.A., Swindle, T.D., and Kring, D.A. (2000). Support for the Lunar Cataclysm Hypothesis from Lunar Meteorite Impact Melt Ages. Science *290*, 1754-1756.

47. Drummond, A.J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol *7*, 214.

48. Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Gateway Computing Environments Workshop (GCE), 2010. pp. 1-8.

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

50. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res *41*, D590-D596.

51. Nawrocki, E.P., Kolbe, D.L., and Eddy, S.R. (2009). Infernal 1.0: inference of RNA alignments. Bioinformatics *25*, 1335-1337.

52. Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., and Tiedje, J.M. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res *42*, D633-D642.

53. Letunic, I., and Bork, P. (2011). Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res *39*, W475-W478.

54. Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics.

55. Hofmann, H.J. (1976). Precambrian Microflora, Belcher Islands, Canada: Significance and Systematics. J Paleontol *50*, 1040-1073.

56. Fischer, W.W. (2008). Biogeochemistry: Life before the rise of oxygen. Nature *455*, 1051-1052.

57. Shih, Patrick M. (2015). Cyanobacterial Evolution: Fresh Insight into Ancient Questions. Current Biology *25*, R192-R193.

58. Nakayama, T., Kamikawa, R., Tanifuji, G., Kashiyama, Y., Ohkouchi, N., Archibald, J.M., and Inagaki, Y. (2014). Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. Proc Natl Acad Sci *111*, 11407-11412.

59. Tripp, H.J., Bench, S.R., Turk, K.A., Foster, R.A., Desany, B.A., Niazi, F., Affourtit, J.P., and Zehr, J.P. (2010). Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. Nature *464*, 90-94.

60. Pereira, M.M., Santana, M., and Teixeira, M. (2001). A novel scenario for the evolution of haem–copper oxygen reductases. Biochim Biophys Acta *1505*, 185-208.

61. Johnson, J.E., Webb, S.M., Thomas, K., Ono, S., Kirschvink, J.L., and Fischer, W.W. (2013). Manganese-oxidizing photosynthesis before the rise of cyanobacteria. Proc Natl Acad Sci.

62. Anbar, A.D., Duan, Y., Lyons, T.W., Arnold, G.L., Kendall, B., Creaser, R.A., Kaufman, A.J., Gordon, G.W., Scott, C., Garvin, J., et al. (2007). A Whiff of Oxygen Before the Great Oxidation Event? Science *317*, 1903-1906.

63. Godfrey, L.V., and Falkowski, P.G. (2009). The cycling and redox state of nitrogen in the Archaean ocean. Nature Geosci *2*, 725-729.

64. Liang, M.-C., Hartman, H., Kopp, R.E., Kirschvink, J.L., and Yung, Y.L. (2006). Production of hydrogen peroxide in the atmosphere of a Snowball Earth and the origin of oxygenic photosynthesis. Proc Natl Acad Sci *103*, 18896-18899.

65. Ettwig, K.F., Butler, M.K., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M.M.M., Schreiber, F., Dutilh, B.E., Zedelius, J., de Beer, D., et al. (2010). Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature *464*, 543-548.

66. Mulkidjanian, A.Y., Koonin, E.V., Makarova, K.S., Mekhedov, S.L., Sorokin, A., Wolf, Y.I., Dufresne, A., Partensky, F., Burd, H., Kaznadzey, D., et al. (2006). The cyanobacterial genome core and the origin of photosynthesis. Proc Natl Acad Sci *103*, 13126-13131.

67. Raymond, J., Zhaxybayeva, O., Gogarten, J.P., Gerdes, S.Y., and Blankenship, R.E. (2002). Whole-Genome Analysis of Photosynthetic Prokaryotes. Science *298*, 1616-1620.

68. Hohmann-Marriott, M.F., and Blankenship, R.E. (2011). Evolution of Photosynthesis. Annu Rev Plant Biol *62*, 515-548.

69. Tice, M.M., and Lowe, D.R. (2004). Photosynthetic microbial mats in the 3,416-Myr-old ocean. Nature *431*, 549-552.

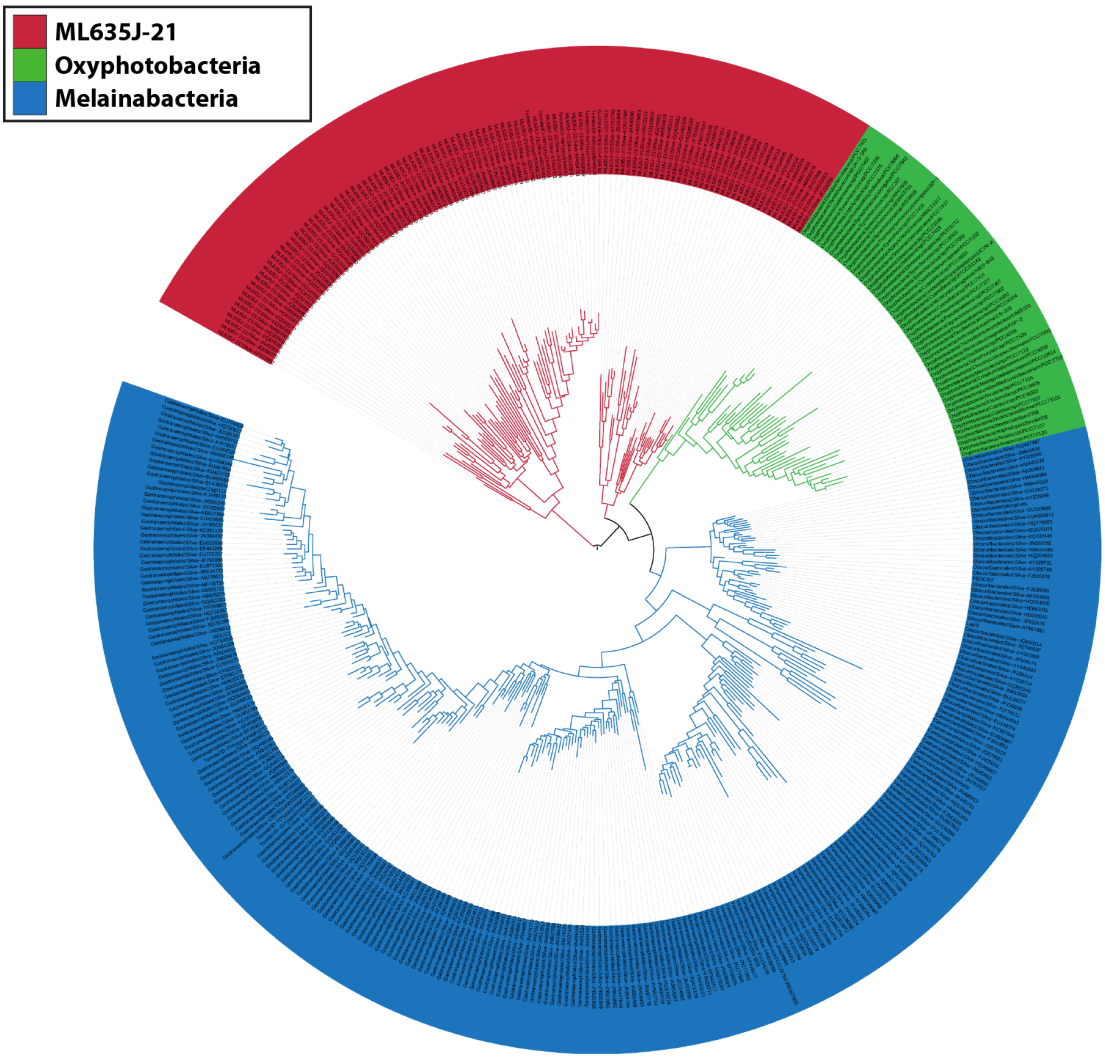


Figure 1. Complete labeled 16S rDNA tree of the Cyanobacterial phylum. Class Oxyphotobacteria in green, class Melainabacteria (orders Gastranaerophilales, Obscuribacterales, Vampirovibrionales, Caenarcaniphilales) in blue, and ML635J-21 clades shown in red. The majority of these Melainabacteria and ML635J-21 sequences were collected from aphotic and anaerobic environments, consistent with the physiologies reconstructed from existing Melainabacteria genomes. This topology warrants consideration of the hypothesis that oxygenic photosynthesis was derived from within the Cyanobacterial phylum.

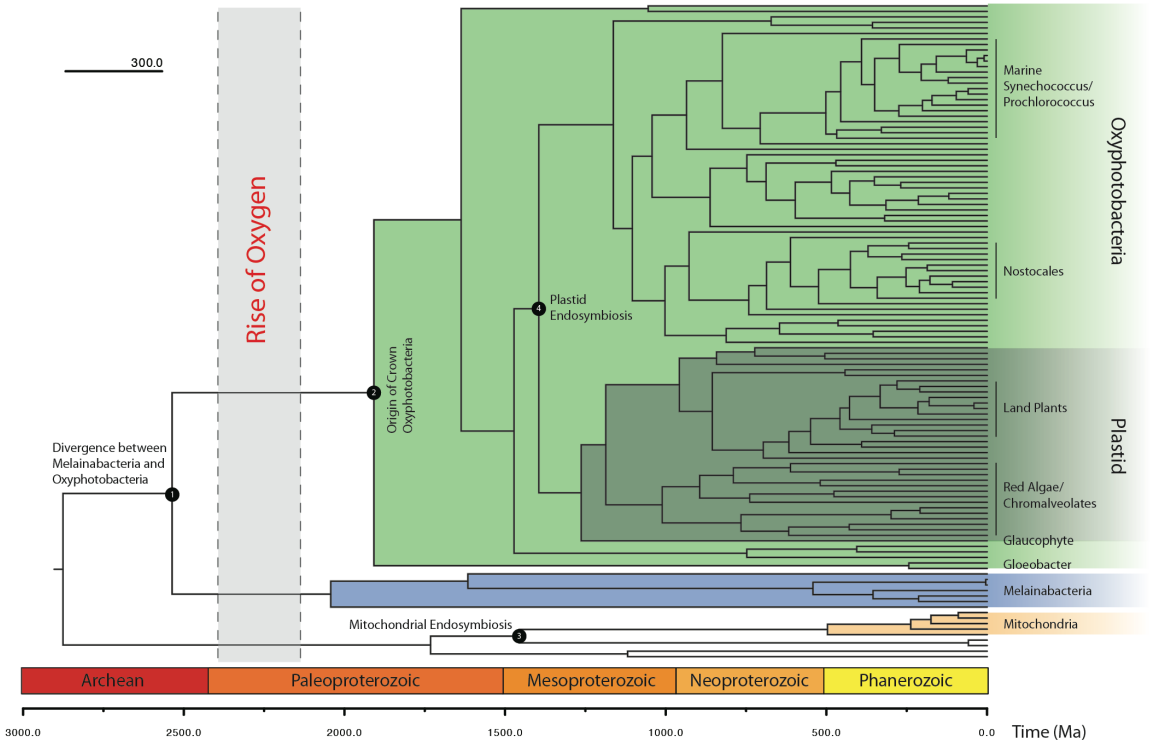


Figure 2. Divergence time estimates for the Cyanobacterial phylum. Dated phylogeny generated from cross-calibrated Bayesian analysis of a concatenated dataset (Run T65). The alignment is composed of conserved proteins found in plastids, mitochondria, and bacteria as well as their 16S rDNA sequences. All analyses illustrate that crown group Oxyphotobacteria postdate the rise of oxygen, which in turn reflects O_2 sourced from stem group lineages. Error bars on ages of nodes are shown and summarized in Figure S2.

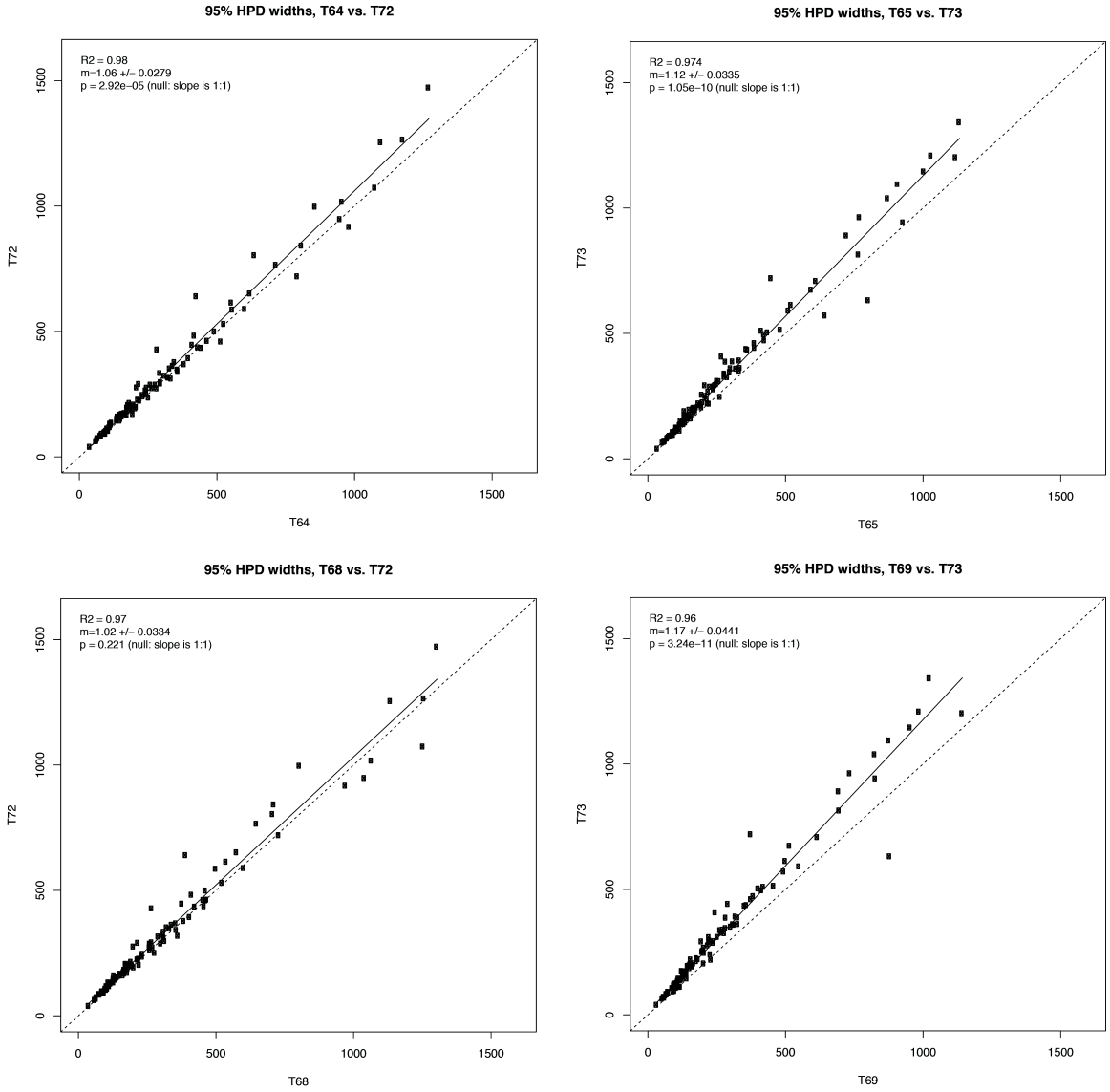


Figure 3. Higher observed age uncertainty when the "Rise of Oxygen" constraint is placed on the radiation of crown group Oxyphotobacteria. Each plot is a comparison of two BEAST analyses, where the width of the 95% HPD represents the amount of dating uncertainty. Each dot represents a corresponding node-date estimate in both trees. All plots are in comparison to one of the two analyses that place the "Rise of Oxygen" constraint on the radiation of crown group Oxyphotobacteria (T72 and T73). In all plots, the observed trend shows that there are higher levels of uncertainty for the T72 and T73 analyses,

suggesting that the placement of this constraint is in poorer agreement with the dataset than the combinations of constraints placed with analyses T64, T65, T68, and T69. *p*-values were calculated using a 1:1 slope as the null hypothesis, assuming that there is no difference in age uncertainty between both analyses.

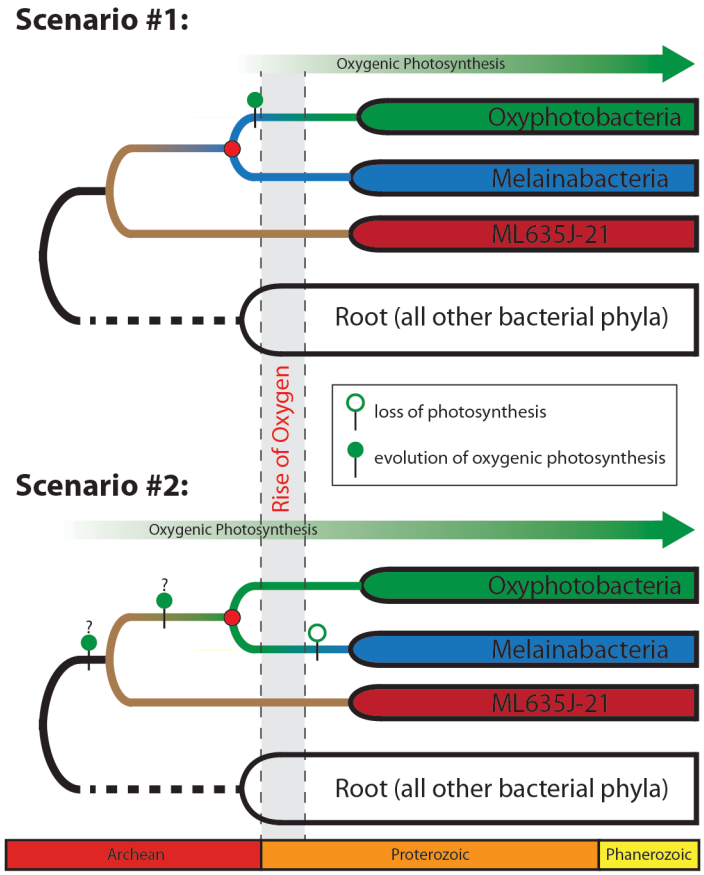


Figure 4. Schematic of two possible scenarios for the timing of oxygenic photosynthesis in relation to the divergence of Oxyphotobacteria and Melainabacteria. Scenario 1. Stem group Oxyphotobacteria diverge from their common ancestor with Melainabacteria (red circle) and evolve oxygenic photosynthesis later. This is the most parsimonious scenario given current data. In this scenario, the divergence of Oxyphotobacteria and Melainabacteria forms a maximum age constraint for oxygenic photosynthesis *ca*. 2.5-2.6 Ga., and is consistent with a wide range of geological and geochemical data for the rise of oxygen at *ca*. 2.35 Ga. **Scenario 2.** Oxygenic photosynthesis evolved prior to the divergence of the Oxyphotobacteria and

Melainabacteria, with subsequent loss(es) in the Melainabacteria. Genomic data from members of ML635J-21 clades will help resolve this ambiguity.

Table 1. Summary of calibration constraints used in this study.

Divergence event	Type of Distribution	Constraint in Mya (±std dev)
Monocotyledoneae	Normal	156 (±14)
Angiospermae	Normal	217 (±40)
Gymnospermatophyta	Normal	327 (±30)
Tracheophyta	Normal	432 (±30)
Land Plants	Normal	477 (±70)
Bangiomorpha	Uniform	1174-1222
"Rise of Oxygen"	Uniform	2400-3000
Last Common Ancestor	Uniform	2400-3800

Table 2. Summary of cross-calibrated BEAST runs generated in this study.

Name	Description	Generations
T64	"Rise of Oxygen" prior set on Mel/Oxyph divergence. Bangiomorpha prior used.	78.04 million post-burnin generations across 5 runs (each run had a burnin of 5 million), 103.04 million total
T65	"Rise of Oxygen" prior set on Mel/Oxyph divergence. Bangiomorpha prior omitted.	79.08 million post-burnin generations across 5 runs (each run had a burnin of 5 million), 104.08 million total
T68	"Rise of Oxygen" prior omitted. Bangiomorpha prior used.	99.12 million post-burnin generations across 5 runs (each run had a burnin of 5 million), 124.12 million total
T69	"Rise of Oxygen" prior omitted. Bangiomorpha prior omitted.	133.14 million post-burnin generations across 5 runs (each run had a burnin of 5 million), 158.14 million total
T72	"Rise of Oxygen" prior set on crown cyano divergence. Bangiomorpha prior used.	81.14 million post-burnin generations using a burnin of 5 million, 106.14 million total
T73	"Rise of Oxygen" prior set on crown cyano divergence. Bangiomorpha prior omitted.	81.78 million post-burnin generations using a burnin of 5 million, 106.78 million total

Table 3. Age estimates for key evolutionary divergences within the oxygenic Cyanobacteria using cross-calibrated methods. Despite substantially different combinations of deep time constraints ("Rise of Oxygen" and "*Bangiomorpha*" constraints), all BEAST runs robustly estimate a Neoarchean divergence between oxygenic Cyanobacteria and Oxyphotobacteria, as well as a Paleoproterozoic radiation of crown group Oxyphotobacteria. Each column identifies the age constraints that were used in different model runs in the study. Dates are in units of millions of years, and parentheses denote the limits of the 95% highest posterior density.

Run	**"Rise of Oxygen" constraint**	**Bangiomorpha constraint**	**Melainabacteria/ Oxyphotobacteria split**	**Crown Oxyphotobacteria**
T64	yes	yes	2630 (2400-2930)	2071 (1805-2393)
T65	yes	no	2536 (2400-2853)	1909 (1556-2254)
T68	no	yes	2541 (2090-3066)	2024 (1723-2374)
T69	no	no	2238 (1750-2790)	1741 (1361-2161)

A p p e n d i x 6

AEROBIC ENVIRONMENTS ON MARS

Vlada Stamenković, Lewis M. Ward, Michael Mischna, & Woodward W. Fischer. "Aerobic Environments on Mars." Nature, submitted.

Aerobic respiration is the most widespread and energetically favorable metabolism on Earth; it enabled complex multicellularity[1]. The history of O_2 and aerobic respiration on the Earth is tied to that of photosynthesis[2], resulting today in ~ 21 vol.% O_2 in an ~1 bar atmosphere. In contrast, modern Mars has a thin, ~6 mbar atmosphere with O_2 produced photochemically from CO_2 and H_2O, and hydrogen escape[3]. Despite the thin atmosphere and overall cold climate, liquid water and brines can exist at and near the surface[4-6]. In particular, Ca- and Mg-perchlorate brines can remain a "normal" liquid above ~200 K and a supercooled liquid down to 140-150 K[7,8]. Using new solutions that describe the solubility of molecular oxygen (O_2) in these and other liquids, we show that, today, all of Mars can host liquid surface environments with sufficient dissolved O_2, $[O_2]_{aq}$, to support aerobic microorganisms, with concentrations that vary substantially with location. Closer to the equator, intermediate $[O_2]_{aq}$ values of 10^{-4}-10^{-3} mol m^{-3} can exist. The highest $[O_2]_{aq}$ values of 10^{-3}-10^{-1} O_2 mol m^{-3} are feasible in cold polar regions—these values are high enough to support respiration by some multicellular basal animals[9]. Such aerobic oases cover ~6.5% of the Martian surface today, but decrease at higher obliquity, or axial tilt, becoming unlikely above ~50°. Our results illustrate that today, and for the

geologically recent past, Mars could have had environments with sufficient levels of dissolved O_2 to support aerobic biology, were it present.

The geologic record of Mars provides evidence of ancient aqueous environments on or near the surface and evidence of great environmental change[10] likely driven by variations in obliquity[11] and atmospheric loss processes[12]. While these observations support the idea that Mars was habitable[13], we don't know if it was ever inhabited; and if so, what metabolisms might have been, or might still be, available to support life.

Recent evidence demonstrates hydrated magnesium and calcium perchlorate salts at various locations on Mars[5,6,14], which indicate the existence of $Mg(ClO_4)_2$-$Ca(ClO_4)_2$-H_2O brines, and are, in some cases, associated with seasonal flow structures like the modern recurring slope lineae[4,5] (RSL). Ca- and Mg-perchlorate brines exhibit a much lower freezing point than pure water, by 60-80 K[7,8]; and, for the eutectic composition, they effectively supercool down to 140-150 K before transitioning into a glass, even when mixed with Martian regolith simulant[7]. Also, theoretical calculations[15] and experiments[7] using brines containing Martian soils with perchlorates (e.g., Phoenix soil analog[15]) show that dissolved perchlorate concentrations can readily reach eutectic concentrations during freezing.

Little attention has been given to the possibility of aerobic respiration on Mars due to the scarcity of O_2 in the Martian atmosphere[3,16]. However, geochemical evidence from Martian meteorites[17] and manganese-rich rocks[18-20] point to highly oxidizing aqueous environments on Mars in its past, suggesting that O_2 played a role in chemical weathering

of Mars crust. Meanwhile, recent experiments, observation, and calculations have lowered the known $[O_2]_{aq}$ limits for aerobic respiration to ~10^{-6} mol m^{-3} in microorganisms[21] and ~$2{\cdot}10^{-3}$ mol m^{-3} in sponges[9]. Understanding the solubility of O_2 in liquid brines, especially containing Ca- and Mg-perchlorates, under Martian conditions is needed to evaluate the spatial and temporal distribution of aerobic environments on Mars.

To achieve this, we developed a new comprehensive thermodynamic framework applicable to Martian conditions that calculates the solubility of O_2 in liquid brines composed of water and salts, such as perchlorates ($Ca(ClO_4)_2$, $Mg(ClO_4)_2$, $NaClO_4$), chlorides ($MgCl_2$, NaCl), and sulfates ($MgSO_4$). We coupled this solubility framework to a Mars general circulation model (GCM) to compute the solubility of O_2 as a function of annually averaged values of pressure and temperature varying with location on the Martian surface today (for obliquity ~25°) and for other values of the planet's obliquity[22] (10°, 15°, 20°, 40°, 60°, 90°). This gives us 3D maps of oxygen solubilities – representing the local, climatically-induced potential for the formation of oxygen-rich brines on Mars for different obliquities. Finally, we used calculations of Mars' obliquity changes in the past ~20 Ma and the next ~10 Myr[23] as input for the GCM to infer how O_2 solubility in surface environments varies across time.

The fundamental challenge was to describe gas dissolution processes in brines below the freezing point of water. To calculate the solubility of O_2 in these liquids, we first computed how the solubility of O_2 changes as a function of temperature in supercooled water. The only uncertainty for this computation concerns the specific heat of dissolved O_2 in supercooled water. For this, we compared a conservative estimate assuming a constant

value (BE, best estimate, in Figures 1-3), tested plausible alternatives resulting from the behavior of the specific heat of supercooled water[24] and brines[24,25] (which led to no significant differences from BE), and derived an unreachable thermodynamic lower limit for the solubility of O_2 (worst case, WC, in Figures 1-3). The presence of dissolved cations and anions (for $Ca(ClO_4)_2$, Ca^{2+} and ClO_4^-) reduces the solubility of O_2 in water[26]. To account for this "salting out"[27] of O_2, we applied the Pitzer approach[28], incorporating O_2-cation, O_2-anion, and O_2-cation-anion electrostatic interactions, and estimated their temperature dependence with the apparent volume model[29] and available experimental data. The BE model results agree well (generally within a few percent) with existing data.

We examined the solubility of O_2 for various potential Martian brines as a function of temperature from 140-300 K for an average Mars surface pressure of ~6.1 mbar, assuming, for each temperature, the critical minimum concentration of salt needed for the brine to remain liquid (Figure 1a). Only Mg- and Ca-perchlorates have been experimentally shown to remain liquid in a supercooled state far below their eutectic temperature to at least ~140-150 K[7] (Figure 1b), and, hence, only for those brines do solubilities increase by at least five orders of magnitude from 273-140 K for BE (Figure 1c). For WC, the O_2 solubilities at 140 K for Ca- and Mg-perchlorates are reduced in comparison to BE but the O_2 solubility still increases by about two orders of magnitude from 273-140 K. All other salts can be only slightly supercooled below their eutectic temperature (Figure 1b), leading to solubilities of O_2 varying by less than one order of magnitude for Martian temperature and average pressure conditions. Over all the solution space, dissolved O_2 concentrations

are in excess of—and commonly orders of magnitude higher than ~10^{-6} mol m^{-3} (Figure 1c, d).

O_2 solubility is primarily controlled by temperature but also by pressure—both of which vary substantially across Mars' surface. To assess spatial variation in dissolved O_2 concentrations, we connected our O_2 solubility model for Ca- and Mg-perchlorate brines with a climate model for modern-day Mars, which accounts for a perennial south polar cap of CO_2 (SPC)[22]. For each temperature, we computed the critical concentration of $Ca(ClO_4)_2$ or $Mg(ClO_4)_2$ needed for the brine to remain liquid, and then solved for $[O_2]_{aq}$ under scenarios with and without supercooling of these brines. We observed large gradients in $[O_2]_{aq}$ values across the Martian surface, with polar regions having the potential to harbor fluids that contain up to 0.2 mol m^{-3} O_2, and the least O_2-rich environments at about ~$2.5 \cdot 10^{-5}$ mol m^{-3} O_2. Without supercooling, liquid perchlorate brines should occur primarily at latitudes below 50°, leading to differences in potential solubility of ~1 order of magnitude depending on location in this belt (see Figure 2b). Even for WC (with and without supercooling), values for the solubility of O_2 in Ca- and Mg-perchlorate brines are above ~10^{-5} mol m^{-3} and vary across the surface by 1-2 orders of magnitude (see Figure 2d).

We extended our results to examine how O_2 solubility in Ca- and Mg-perchlorate brines varies with climates at different obliquities. For BE and WC with supercooling, we observed that the range and average of O_2 solubility significantly decreased for higher obliquities, due to an increase of the minimum values and a decrease of the maximum values for the annually averaged surface temperature across the planet's surface. Integrating these results with obliquity variation through time[22], we found that the past 5

Ma have supported particularly O_2-rich environments (Figure 3b), while the preceding 15 Ma supported average maximum O_2 solubility values ~200 times lower than today. These results demonstrate substantial secular variation in the potential O_2 content of fluids in Mars surface environments over recent geological time.

We found that on modern-day Mars—including all uncertainties, for our best estimate and worst case, with and without supercooling—the solubility of O_2 in various fluids exceeds by at least 1-2 orders of magnitude the level required for aerobic respiration[21] by microbes (see Figure 1e and 2). Thus, Mars could offer a wide range of surface and near-surface environments containing enough dissolved oxygen to support aerobic respiration like that seen in diverse groups of terrestrial microorganisms. Moreover, for supercooled Ca- and Mg-perchlorate brines on Mars today, ~6.5% of the surface at latitudes above 70° (north and south) could support dissolved O_2 concentrations far higher—levels sufficient to sustain respiration demands of complex multicellular organisms (Figure 1d). Such aerobic oases with $[O_2]_{aq}$>$2 \cdot 10^{-3}$ mol m^{-3} are feasible for low obliquities and become unlikely for obliquities above ~50° (Figure 3a) (such high-obliquity states did not happen in the last 20 Ma, but might have existed in the deeper Martian past[23]). Other aerobic environments with $[O_2]_{aq}$ ~10^{-4}-10^{-3} mol m^{-3} can occur closer to the equator in higher pressure regions like Hellas, Arabia Terra, Amazonis Planitia, and Tempe Terra (Figure 2).

At the levels for molecular oxygen solubility that we find, O_2 would be expected to play a role in chemical weathering of Mars' crust[30]—potentially explaining a range of geochemical and geologic data, such as highly oxidized phases in Martian rocks[18,19].

We do not know whether Mars was ever a host to life. On Earth, aerobic respiration appears to have followed in the evolutionary footsteps of oxygenic photosynthesis. However, Mars shows us this need not be the case, broadening our view of the opportunities for aerobes on other planetary bodies.

1 Figures for main text

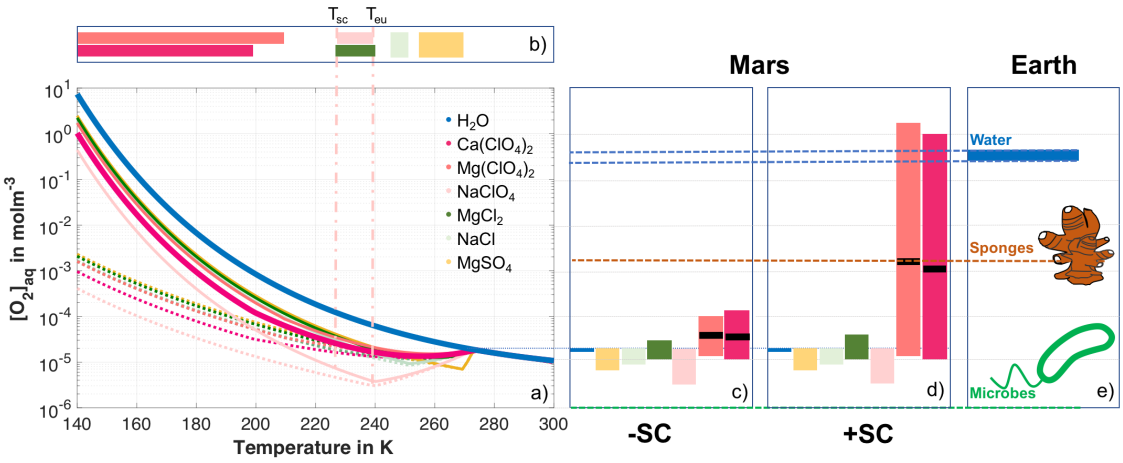


Figure 1 – Solubility of O_2, $[O_2]_{aq}$, in liquid water and brines on Mars. a) $[O_2]_{aq}$ change with temperature at a pressure of 6.1 mbar using the minimum salt concentration that keeps the brine liquid[7,8,25] (best estimate, BE, is *solid*, worst case, WC, is *dashed)*. **b)** Supercooling feasible between eutectic temperature, T_{eu}, and T_{sc}[7,8,25]. **c, d)** Bars represent BE $[O_2]_{aq}$ range with (+SC) and without supercooling (-SC). S*olid black* line shows upper $[O_2]_{aq}$ limit reduction (if any) for WC. **e)** Blue bar shows the $[O_2]_{aq}$ range for water with *T*~273-300 K and 1 bar, dashed lines show the O_2 levels required for aerobic respiration by microbes[21] (green) and animals[9] (orange).

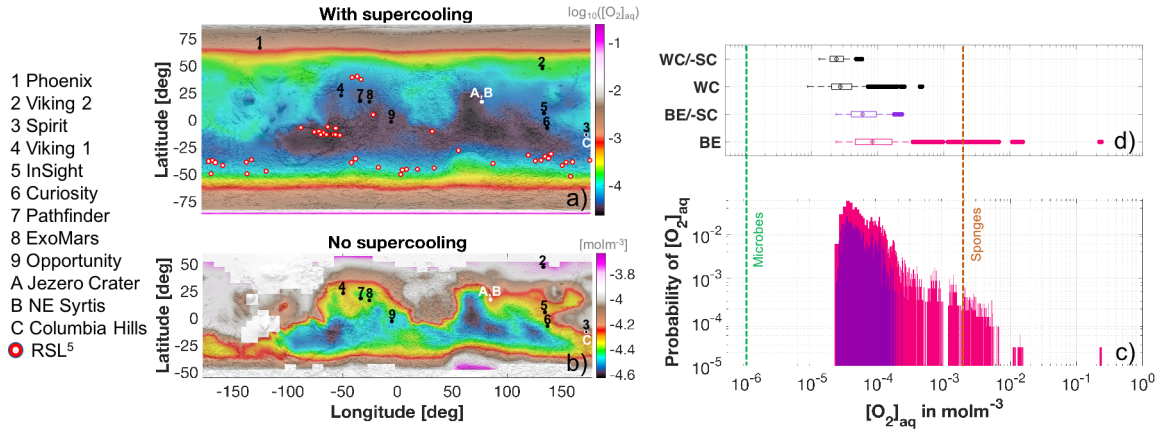


Figure 2 – Climatically-induced spatial variation of O_2 solubility in Ca-perchlorate brines on modern-day Mars: a, b) Maps of $[O_2]_{aq}$ according to the best estimate (BE) for the annually averaged O_2 solubility $[O_2]_{aq}$ in Ca-perchlorate brines with (a) and without (b) supercooling under modern Martian conditions with a perennial south polar CO_2 cap. Without supercooling, only latitudes within ~50° favor liquid brines. **c)** Histogram of $[O_2]_{aq}$ values for BE with (red) and without (purple) supercooling and biological breathing limits. **d)** Boxplots of BE and worst case (WC) with supercooling and without (/-SC). $[O_2]_{aq}$ exceeds the lower limit for aerobic respiration by microbes, and in certain environments, surpasses the $[O_2]_{aq}$ limit for sponges. Results very similar for Mg-perchlorate.

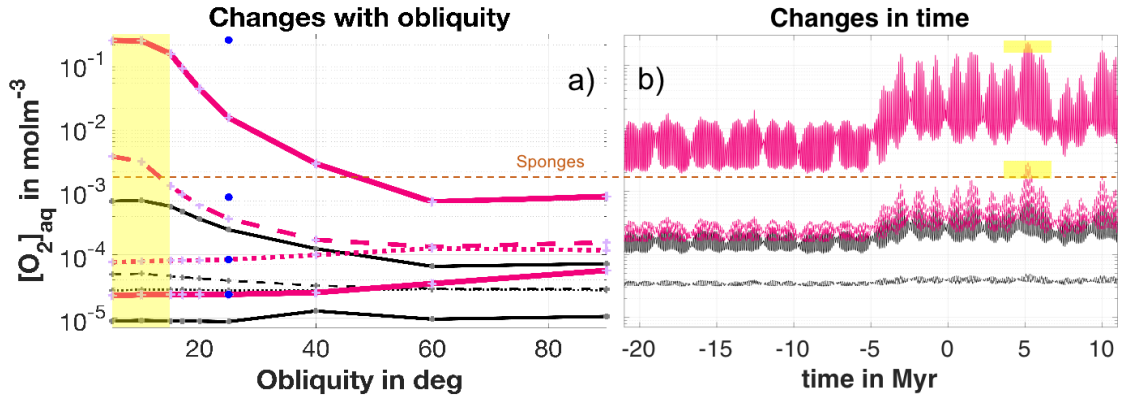


Figure 3 – Obliquity-driven evolution of aerobic environments on Mars across time: **a)** With supercooling, best estimate (BE, purple crosses with red line) and worst case (WC, gray crosses with black line) for global maximum (top solid), minimum (bottom solid), average (dashed), and median (dotted) O_2 solubilities in Ca-perchlorate brines with varying obliquity without a perennial south polar cap (SPC) (blue dots are BE results with SPC, see Figure 2). **b)** Calculations of how $[O_2]_{aq}$ varies during past and future obliquity changes[23] showing the global maximum (top) and average (bottom) $[O_2]_{aq}$ values across time. Highlighted in a dashed orange line is the aerobic limit for sponges and obliquities/times where atmospheric collapse occurs (yellow).

References

1 Catling, D. C., Glein, C. R., Zahnle, K. J. & McKay, C. P. Why O2 Is Required by Complex Life on Habitable Planets and the Concept of Planetary "Oxygenation Time". *Astrobiology* **5**, 415-438 (2005).

2 Fischer, W. W., Hemp, J. & Johnson, J. E. Evolution of Oxygenic Photosynthesis. *Annu. Rev. Earth Planet. Sci.* **44**, 647–683, doi:10.1146/annurev-earth-060313-054810 (2016).

3 Nair, H., Allen, M., Anbar, A. D., Yung, Y. L. & Clancy, R. T. A Photochemical Model of the Martian Atmosphere. *Icarus* **111**, 124-150, doi:10.1006/icar.1994.1137 (1994).

4 Ojha, L., et al. Spectral evidence for hydrated salts in recurring slope lineae on Mars. *Nature Geoscience* **8**, 829-832, doi:10.1038/ngeo2546 (2015).

5 Rummel, J. D. *et al*. A new analysis of Mars "Special Regions": findings of the second MEPAG Special Regions Science Analysis Group (SR-SAG2). *Astrobiology* **14**, 887-968, doi:10.1089/ast.2014.1227 (2014).

6 Kounaves, S. P. *et al*. Identification of the perchlorate parent salts at the Phoenix Mars landing site and possible implications. *Icarus* **232**, 226–231, doi:10.1016/j.icarus.2014.01.016 (2014).

7 Toner, J. D., Catling, D. C. & Light, B. The formation of supercooled brines, viscous liquids, and low-temperature perchlorate glasses in aqueous solutions relevant to Mars. *Icarus* **233**, 36-47, doi:10.1016/j.icarus.2014.01.018 (2014).

8 Pestova, O. N., Myund, L. A., Khripun, M. K. & Prigaro, A. V. Polythermal Study of the Systems M(ClO4)2-H2O (M2+ = Mg2+, Ca2+, Sr2+, Ba2+). *Russian Journal of Applied Chemistry* **78**, 409-413 (2005).

9 Mills, D. B. *et al*. Oxygen requirements of the earliest animals. *Proc Natl Acad Sci U S A* **111**, 4168-4172, doi:10.1073/pnas.1400547111 (2014).

10 Grotzinger, J. P. & Milliken, R. E. The sedimentary record of Mars: distribution, origins, and global stratigraphy. *Sedimentary Geology of Mars,* SEPM Special Publication, 1-48 (2012).

11 François, L. M., Walker, J. C. G. & Kuhn, W. R. A numerical simulation of climate changes during the obliquity cycle on Mars. *Journal of Geophysical Research* **95**, 14761, doi:10.1029/JB095iB09p14761 (1990).

12 Jakosky, B. M. *et al*. Mars' atmospheric history derived from upper-atmosphere measurements of 38Ar/36Ar. *Science* **355**, 1408–1410 (2017).

13 Grotzinger, J. P. *et al*. Deposition, exhumation, and paleoclimate of an ancient lake deposit, Gale crater, Mars. *Science* **350**, aac7575, doi:10.1126/science.aac7575 (2015).

14 Leshin, L. A. *et al*. Volatile, isotope, and organic analysis of martian fines with the Mars Curiosity rover. *Science* **341**, 1238937, doi:10.1126/science.1238937 (2013).

15 Marion, G. M., Catling, D. C., Zahnle, K. J. & Claire, M. W. Modeling aqueous perchlorate chemistries with applications to Mars. *Icarus* **207**, 675-685, doi:10.1016/j.icarus.2009.12.003 (2010).

16 Mahaffy, P. R. *et al*. Abundance and isotopic composition of gases in the martian atmosphere from the Curiosity rover. *Science* **341**, 263-266, doi:10.1126/science.1237966 (2013).

17 Shaheen, R., Niles, P. B., Chong, K., Corrigan, C. M. & Thiemens, M. H. Carbonate formation events in ALH 84001 trace the evolution of the Martian atmosphere. *Proc Natl Acad Sci U S A* **112**, 336-341, doi:10.1073/pnas.1315615112 (2015).

18 Lanza, N. L., et al. . High manganese concentrations in rocks at Gale crater, Mars. *Geophys. Res. Lett.* **41**, 5755–5763, doi:10.1002/2014GL060329 (2014).

19 Lanza, N. L., et al. . Oxidation of manganese in an ancient aquifer, Kimberley formation, Gale crater, Mars. *Geophys. Res. Lett.* **43**, 7398–7407, doi:10.1002/2016GL069109 (2016).

20 Arvidson, R. E. *et al*. High concentrations of manganese and sulfur in deposits on Murray Ridge, Endeavour Crater, Mars. *American Mineralogist* **101**, 1389–1405, doi:10.2138/am-2016-5599 (2016).

21 Zakem, E. J. & Follows, M. J. A theoretical basis for a nanomolar critical oxygen concentration. *Limnology and Oceanography* **62**, 795-805, doi:10.1002/lno.10461 (2017).

22 Mischna, M. A., Baker, V., Milliken, R., Richardson, M. & Lee, C. Effects of obliquity and water vapor/trace gas greenhouses in the early martian climate. *Journal of Geophysical Research: Planets* **118**, 560–576, doi:10.1002/jgre.20054 (2013).

23 Laskar, J. *et al*. Long term evolution and chaotic diffusion of the insolation quantities of Mars. *Icarus* **170**, 343-364, doi:10.1016/j.icarus.2004.04.005 (2004).

24 Archer, D. G. & Carter, R. W. Thermodynamic Properties of the NaCl + H2O System. 4. Heat Capacities of H2O and NaCl(aq) in Cold-Stable and Supercooled States. *J. Phys. Chem. B* **104**, 8563-8584 (2000).

25 Toner, J. D. & Catling, D. C. Water activities of NaClO4, Ca(ClO4)2, and Mg(ClO4)2 brines from experimental heat capacities: Water activity >0.6 below 200K. *Geochimica et Cosmochimica Acta* **181**, 164-174, doi:10.1016/j.gca.2016.03.005 (2016).

26 Clegg, S. L. & Brimblecombe, P. The solubility and activity coefficient of oxygen in salt solutions and brines. *Geochimica et Cosmochimica Acta* **54**, 3315-3328 (1990).

27 Konnik, E. I. Salting-out and Salting-in of Gaseous Non-electrolytes in Aqueous Solutions of Electrolytes. *Russian Chemical reviews* **46**, 577-588 (1977).

28 Pitzer, K. S. Theoretical considerations of solubility with emphasis on mixed aqueous electrolytes. *Pure & AppI. Chem.* **58**, 1599-1610 (1989).

29 Tromans, D. Modeling Oxygen Solubility in Water and Electrolyte Solutions. *Ind. Eng. Chem. Res.* **39**, 805-812 (2000).

30 Johnson, J. E., Gerpheide, A., Lamb, M. P. & Fischer, W. W. O2 constraints from Paleoproterozoic detrital pyrite and uraninite. *Geological Society of America Bulletin*, doi:10.1130/B30949.1 (2014).

2 Methods

Below we summarized all methods—from solubility to global circulation model—needed to reproduce our results. We direct the reader to our Supplementary Online Material (SOM) for details on any of the methods subsections.

Solubility model. The solubility of O_2 is the number of dissolved O_2 molecules ($[\widetilde{O_2}]_{aq,X}$ in mol kg^{-1} and $[O_2]_{aq,X}$ in mol m^{-3}) in a brine consisting of m_X mol of salt X per kg of pure water in equilibrium with an atmosphere of pressure, P, partial O_2 volume fraction, f_{O_2} ($f_{O_2}P$ corresponds to the total partial O_2 pressure in the atmosphere), and surface temperature, T.

To compute the solubility, $[\widetilde{O_2}]_{aq,X}$, we started with the solubility of O_2 in pure water, $[\widetilde{O_2}]_{aq,w}$.

$$[\widetilde{O_2}]_{aq,w}(T,P) = \frac{f_{O_2}P}{P_{ref}} exp\left(\frac{\tilde{\mu}_{g,O_2}(T) - \tilde{\mu}_{aq,w,O_2}(T)}{RT}\right) \quad (1)$$

$\tilde{\mu}(T)$ is the chemical potential as a function of temperature at fixed reference pressure P_{ref} – with $\tilde{\mu}_{g,O_2}(T)$ for O_2 in the gas phase and $\tilde{\mu}_{aq,w,O_2}$ for O_2 dissolved in water.

To compute the solubility of O_2 in pure water, $[\widetilde{O_2}]_{aq,w}$, we needed to determine the difference in molar chemical potential evaluated at the reference pressure P_{ref} between O_2 dissolved in pure water and in the gas phase $\Delta\tilde{G}(T) = \tilde{\mu}_{aq,w,O_2}(T) - \tilde{\mu}_{g,O_2}(T)$as a function of temperature, in J mol^{-1}. The difference in chemical potential corresponds to the difference in molar Gibbs potential between O_2 in the gaseous and dissolved phases.

$$\tilde{\mu}(T) = \tilde{\mu}(T_{ref}) - S(T_{ref})[T - T_{ref}] + \int_{T_{ref}}^{T} C_P(T')dT' - T\int_{T_{ref}}^{T} \frac{C_P(T')}{T'}dT' \quad (2)$$

To calculate the chemical potential as a function of temperature at fixed reference pressure, $\tilde{\mu}(T)$, either for O_2 in the gas phase or in the dissolved phase, we needed to know the values of their chemical potentials $\tilde{\mu}(T_{ref})$ and molar entropy $S(T_{ref})$, in J K^{-1} mol^{-1}, at a reference temperature, T_{ref}—both were previously measured in experiments[31]—and the temperature dependence of the molar specific heat at constant pressure, $C_P(T)$, in J K^{-1} mol^{-1}. The parameters needed to solve Equations 1-2 are listed in Table S1 and can be found in the literature[31]. For the choice of the temperature dependence of the specific heat capacity at constant pressure of gaseous and dissolved O_2 see the Section below on "*Specific heats and robustness of conclusions*".

After having computed the solubility of O_2 in pure supercooled water, we needed to quantify the electrostatic interactions between the water dipoles and the ions in solution because this causes a reduction in apparent water volume that can be used to dissolve O_2[29,31,32], and hence the solubility of O_2 in a brine containing a salt, X, of molality m_X is reduced by a salting out factor of $\gamma_{O_2}(X, m_X)$ compared to pure water.

$$[\widetilde{O_2}]_{aq,X}(T,P) = \frac{1}{\gamma_{O_2}(X, m_X)}\,[\widetilde{O_2}]_{aq,w}(T,P) \quad (3)$$

Experiments demonstrated that the salting out factor $\gamma_{O_2}(X, m_X)$ depends only weakly on temperature[26,33], and thus the temperature dependence is usually neglected[33,34]. In the SOM, we added all the available data on various salts from 300-240 K as well as theoretical calculations, which indicate that too for perchlorates, we expect from 298-140 K

only a weak temperature-induced increase of $\gamma_{O_2}(X, m_X)$ by less than a factor of 3-10 (and likely less than 5).

The electrostatic interactions can be separated into first-order direct undisturbed interactions between the O_2 molecules and the dissolved cations, (O_2-c), and anions, (O_2-a), respectively. Such interactions are described with the Pitzer approach[28] through interaction coefficients λ_{O_2-c} and λ_{O_2-a} and the corresponding molalities for the cations m_c and anions m_a respectively. The secondary interaction between the triplet of cations, anions, and O_2 molecules λ_{O_2-c-a} is negative[26,28] because secondary interactions generally weaken the electrostatic primary exchange of cations or anions with the O_2 molecule, which therefore decreases the salting out factor, and leads to an effective increase of the gas solubility. Setting $\lambda_{O_2-c-a} = 0$ marks hence a lower bound for the solubility. To make the salting out factor a function of salt molality, m_X, we used for each ion (cation and anion) the mol(ion)/mol(salt) fractions, f_c and f_a. The parameters needed to solve Equations 3-4 can be found in Table S2, and can be taken or computed with available data[26,27,35].

$$\gamma_{O_2}(X) = exp\left(\sum_c 2\lambda_{O_2-c} m_c + \sum_a 2\lambda_{O_2-a} m_a + \sum_c \sum_a \lambda_{O_2-c-a} m_c m_a\right)$$

$$= exp\left(2\sum_c \lambda_{O_2-c} f_c m_X + 2\sum_a \lambda_{O_2-a} f_a m_X + \sum_a \sum_c \lambda_{O_2-c-a} f_a f_c m_X^2\right) \quad (4)$$

We converted the solubility of O_2 from mol kg^{-1} $[\widetilde{O_2}]_{aq,X}$ into mol m^{-3} $[O_2]_{aq,X}$ (which is shown in Figures 1-3) because it is the volume number density that is relevant to chemistry and respiration.

$$[O_2]_{aq,X}(T,P) = \rho_{brine}(T,P) \cdot [\widetilde{O_2}]_{aq,X}(T,P) \quad (5)$$

The density of brines is only slightly temperature-dependent, can be well approximated by an incompressible fluid, and increases with decreasing temperatures by less than 5-10% from 298-140 K (see SOM). Therefore, by making the assumption that $\rho_{brine}(T,P) \approx \rho_w(T_{ref}, P_{ref}) \approx 1000 kgm^{-3}$, we attained a lower estimate for the solubility of O_2.

Together, Equations 1-5 allowed us to calculate the solubility of O_2 in a variety of brines, across a wide range of temperatures, including those far below the freezing point of pure water, with the specific heat of dissolved O_2 as the major modulating factor.

Specific heats and robustness of conclusions. The specific heat at constant pressure of O_2 in the gas phase is experimentally well known down to 100 K and remains approximately constant[36] from 300-100 K with $C_{P,O_2,g}(T) \approx 29.332$ J K^{-1} mol^{-1} due to translational and rotational degrees of freedom remaining active throughout this temperature range.

The major factor impacting the solubility of O_2 in pure water is the heat capacity of dissolved O_2 as a function of temperature $C_{P,O_2,aq}(T)$. Larger values of $C_{P,O_2,aq}(T)$ lead to a greater gas solubility. For temperatures between 273-373 K, experiments showed that $C_{P,O_2,aq}(T)$ increases linearly with decreasing temperature towards 273 K, there is no data below 273 K for $C_{P,O_2,aq}(T)$. As we demonstrated in the SOM, there is good evidence that a trend of increasing heat capacity persists until at least ~225-235 K, where water might have a liquid-liquid phase transition[24], possibly leading to a decay of the specific heat of dissolved O_2 for smaller temperatures.

Because of this uncertainty, we took a more conservative approach by considering 1) a best estimate (BE) with $C_{P,O_2,aq}(T) = C_{P,O_2,aq}(T = 298\ K)$; 2) reasonable alternatives for the specific heat capacity as a function of temperature either based on the specific heat for supercooled water or a linear decay towards 140 K (similar to the specific heat of various brines[24,25]) — leading to results of only less than a factor of ~3 smaller at 140 K than what we found with BE; demonstrating the robustness of our conclusions with BE; and 3) a worst-case limit to the solubility of O_2 in water (WC) (see below, and in much more detail in the Supplementary Online Material), as shown in Figure 1 and Figure S1. WC is technically an unreachable limit that would lead at 140 K to a solubility about three orders of magnitude smaller than found with BE. Although this worst case would limit high-O_2 oases with concentrations above the breathing limit for sponges[9], it would still permit O_2 solubilities orders of magnitude above the aerobic limit for microorganisms[21] for Martian conditions.

Thermodynamic limit and worst-case scenario. The specific heat of dissolved O_2, $C_{P,O_2,aq}$, controls the solubility of O_2 in pure water or brine in the following manner (using Equations 1 and 3; compare with Equations S11 and S13):

$[O_2]_{aq}(T,\zeta) = F(T)exp\left(\frac{\zeta(T)}{RT}\right)$ with

$$\zeta(T) = -\int_{T_{ref}}^{T} C_P(T')dT' + T\int_{T_{ref}}^{T} \frac{C_P(T')}{T'}dT' \quad (6)$$

$F(T)$ is defined as the solubility of O_2 if the specific heat of dissolved O_2 in water were always zero, and is well-constrained and monotonically increases for smaller temperatures.

The function $\zeta(T)$, defined in Equation 6, contains all the uncertainties introduced by the behavior of $C_{P,O_2,aq}(T)$ below 273 K. In the SOM, we demonstrated that the function $\zeta(T)$ must also monotonically increase with decreasing temperatures due to the specific heat being always larger than zero, for $T > 0$ K.

Therefore, we can obtain a thermodynamic limit to the solubility of O_2 in pure water by setting the specific heat of dissolved O_2 for temperatures below 273 K to an unreachable minimum value of zero, and, hence, the minimum solubility of O_2 in pure water $T < 273\ K$ is given by:

$$[O_2]_{aq}(T,\zeta)\big|_{min} = [O_2]_{aq}(T,\zeta^*) \text{ with } \zeta^*(T) = \zeta(T = 273\ K) \quad (7)$$

Melting curve for brines. Values for the critical concentration at temperature, T, needed to keep the brine liquid, $m_X(T)$, were taken from experimental data[7,25,37] and fit with a third-degree polynomial to generate a melting curve for the brine. Inside the temperature domain where the brine is known to effectively supercool (between T_{eu} and T_{sc}, see Figure 1b) we assumed $m_{eu} = m_X(T_{eu})$, as no additional salt has to be added to keep the brine liquid, where T_{eu} and m_{eu} are the eutectic temperature and the eutectic concentration respectively. Therefore, we obtained $m_X(T) = \sum_{i=0}^{3} p_i T^i$ for $T > T_{eu}$ and $m_X(T) = m_{eu} = m_X(T_{eu})$ for $T \leq T_{eu}$. Moreover, the eutectic temperature is not a strong function of pressure, and the pressures on Mars support the liquid state of the Ca-perchlorate brines assumed in Figure 2 (see SOM). For parameters, see Table S3.

Boxplots and statistics. We used the statistical toolbox in MATLAB 2016a. The box

width (shown in Figure 2) represents the inter quartile range, with notch marks at the 95% confidence interval for the median circle. The dashed whiskers indicate 2.698 σ. Data beyond 2.7 σ are show in dots outside the whiskers.

General circulation model for Mars. We used the Mars Weather Research and Forecasting (MarsWRF) GCM for this investigation. MarsWRF[38] is a Mars-specific implementation of the PlanetWRF GCM[39]—a global model derived from the terrestrial mesoscale WRF model[40]. MarsWRF solves the primitive equations using a finite difference methodology on an Arakawa-C grid[41]. Both the horizontal and vertical resolution of the model are variable and selectable at run time; a 40-layer vertical grid (from 0-80 km) was used, following a modified-sigma (terrain-following) coordinate. The lowest model layer is ~75-100 m above ground level, depending on location and season. We used a horizontal resolution of 5°x5°, which corresponded to a grid of 72 points in longitude x 36 points in latitude. The total present-day atmospheric CO_2 budget, as well as the CO_2 ice albedo and emissivity for each hemisphere, were adjusted until the modeled pressure curves best matched those observed at the Viking Lander 1 and 2 sites[42]. Both surface albedo and thermal inertia were matched to MGS-TES observations[43,44], while water ice albedo and emissivity were fixed at 0.45 and 1.0, respectively.

The model includes a basic water cycle, which allows the condensation, sublimation, sedimentation, and transport of water ice particles in the atmosphere, and growth and recession of polar water ice caps. Surface albedo was modified to the aforementioned values when either CO_2 or water ice condensed on the surface.

Planetary obliquity, or axial tilt, can be modified at runtime in MarsWRF[45]. In all obliquity simulations, we held other model parameters constant. To examine the past 20 Ma of Mars history, we assumed the luminosity of the Sun to be at the present-day value. Because the frost-point temperature for CO_2 (~148 K at present surface pressure) was regularly reached in the polar winter hemisphere, the present-day Martian atmosphere cycles ~25% of its total mass over the course of the year into and out of the polar caps. For obliquities higher than present-day (~25°), this fraction is somewhat larger. Furthermore, for obliquities below ~15°, solar heating of the polar regions was insufficient to stave off condensation year-round, and Mars experiences episodes of atmospheric collapse, where the condensable species in the atmosphere (predominantly CO_2) formed thick ice caps at the poles. These intervals are indicated in light green in Figure 3, and, if theoretical predictions of obliquity as a function of time are correct, are not seen in the past 20 Ma and will rarely occur in the next 10 Myr[23].

At higher obliquities, water ice preferentially condensed at lower latitudes. While there is disagreement over exactly how widespread such water deposits might be[45,46], MarsWRF handles the tropical ice deposits self-consistently, increasing surface albedo at locations where ice is deposited. With all else equal, Mars at high obliquity will have a slightly lower global surface temperature than at present values due to the increased surface area in the tropics covered by high albedo ice, as compared to the relatively small polar area covered by ice at present. However, the maximum (minimum) values of annually averaged surface temperature will generally decrease (increase) with increasing obliquity – mainly due to the albedo changes associated with obliquity change as mentioned before (see SOM for a more

detailed discussion). The distribution, minimum, average, and maximum values of annually averaged surface temperature can be found in Table S4 as well as in Figure S4. These simulations showed that obliquity variation with time was the major driver for differences in annual averages for surface temperature and pressure; past variations in eccentricity played only a minor role in affecting annually averaged surface temperatures and pressures.

To compare different obliquities, we do not account for Mars' south polar residual CO_2 cap (SPC), which would only reduce the annually averaged surface temperatures, and hence increase the solubility of O_2, as shown for a modern-day obliquity of ~25° in Figure 3a.

Additional references for methods

31 Tromans, D. Temperature and pressure dependent solubility of oxygen in water: a thermodynamic analysis. *Hydrometallurgy* **48**, 327-342 (1998).

32 Tromans, D. Oxygen solubility modeling in inorganic solutions: concentration, temperature and pressure effects. *Hydrometallurgy* **50**, 279–296 (1998).

33 Silvester, L. F. & Pitzer, K. S. Thermodynamics of electrolytes. X. Enthalpy and the effect of temperature on the activity coefficients. *Journal of Solution Chemistry* **7**, 327–337, doi:10.1007/BF00662893 (1978).

34 Toner, J. D., Catling, D. C. & Light, B. A revised Pitzer model for low-temperature soluble salt assemblages at the Phoenix site, Mars. *Geochimica et Cosmochimica Acta* **166**, 327-343, doi:10.1016/j.gca.2015.06.011 (2015).

35 Khomutov N. E., K. E. I. Solubility of oxygen in aqueous electrolyte solutions. *Russian Journal of Physical Chemistry* **48** (1974).

36 Manion, J. A. *et al*. in *NIST Standard Reference Database 17* Vol. 17, Version 7.0 (Web Version), Release 1.6.8, Data Version 2016.10 (National Institute of Standards and Technology, Gaithersburg, Maryland, 20899-8320, 2016).

37 Li, D. *et al*. Phase diagrams and thermochemical modeling of salt lake brine systems.II. NaCl+H2O, KCl+H2O, MgCl2+H2O and CaCl2+H2O systems. *CALPHAD: Computer Coupling of Phase Diagrams and Thermochemistry* **53**, 78-89, doi:10.1016/j.calphad.2016.03.007 (2016).

38 Toigo, A. D., Lee, C., Newman, C. E. & Richardson, M. I. The impact of resolution on the dynamics of the martian global atmosphere: Varying resolution

studies with the MarsWRF GCM. *Icarus* **221**, 276-288, doi:10.1016/j.icarus.2012.07.020 (2012).

39 Richardson, M. I., Toigo, A. D. & Newman, C. E. PlanetWRF: A general purpose, local to global numerical model for planetary atmospheric and climate dynamics. *Journal of Geophysical Research* **112**, doi:10.1029/2006je002825 (2007).

40 Skamarock, W. C. & Klemp, J. B. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of Computational Physics* **227**, 3465-3485, doi:10.1016/j.jcp.2007.01.037 (2008).

41 Arakawa, A. & Lamb, V. R. Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods Comput. Phys.* **17**, 173-265 (1977).

42 Guo, X., Lawson, W. G., Richardson, M. I. & Toigo, A. Fitting the Viking lander surface pressure cycle with a Mars General Circulation Model. *Journal of Geophysical Research* **114**, doi:10.1029/2008je003302 (2009).

43 Christensen, P. R. *et al*. Mars Global Surveyor Thermal Emission Spectrometer experiment: Investigation description and surface science results. *Journal of Geophysical Research* **106**, 23823-23871 (2001).

44 Putzig, N. & Mellon, M. Apparent thermal inertia and the surface heterogeneity of Mars. *Icarus* **191**, 68-94, doi:10.1016/j.icarus.2007.05.013 (2007).

45 Mischna, M. A. On the orbital forcing of Martian water and CO2cycles: A general circulation model study with simplified volatile schemes. *Journal of Geophysical Research* **108**, doi:10.1029/2003je002051 (2003).

46 Forget, F., Haberle, R. M., Montmessin, F., Levrard, B. & Head, J. W. Formation of Glaciers on Mars by Atmospheric Precipitation at High Obliquity. *Science* **311**, 368-371 (2006).

SOM - Supplementary Online Material

In this supplementary section, we derived all equations used in the main article and gave an extended discussion of our methods. It is organized as follows. First, we derived all needed equations in Section 1. In Section 2, we discussed our solubility (2.1) and climate (2.2) models. In Section 3, we explained why the averaging method we employed for computing O_2 solubilities as a function of the annually averaged local values for surface temperature and pressure is practical and underestimates O_2 solubilities and the difference between minimum and maximum values. In Section 4, we elaborate on how pressure and temperature control solubility, and possibly oxidation, gradients across the Martian surface.

1. Detailed derivation of all necessary equations

To compute the solubility of O_2 in pure water $[\widetilde{O_2}]_{aq,w}(T,P)$ in mol kg^{-1} we started with the equilibrium constant for a given pressure and temperature $k(T,P)$, which is defined for an activity coefficient of dissolved O_2, α, and the fugacity coefficient, γ, of O_2 in the gas phase[1-4] with f_{O_2} being the volume fraction of O_2 in the atmosphere (for Mars, $f_{O_2} =$

0.00146[5]).

$$k(T,P) = \frac{\alpha[\widetilde{O_2}]_{aq,w}(T,P)}{\gamma f_{O_2}} \quad \text{(S1)}$$

Both activity coefficient of dissolved O_2 as well as the fugacity of O_2 in the gas phase are close to one because of the small concentrations of dissolved O_2 and the small pressures that are relevant to our study (independent of temperature, for $[\widetilde{O_2}]_{aq,w} \to 0$ and $P \to 0$, the definitions for activity coefficient and fugacity demand that $\alpha \to 1$ and $\gamma \to 0$ respectively[1]). Therefore, Equation S1 simplifies to:

$$k(T,P) = \frac{[\widetilde{O_2}]_{aq,w}(T,P)}{f_{O_2}} \quad \text{(S2)}$$

To compute the solubility $[\widetilde{O_2}]_{aq,w}(T,P)$, we needed to compute the equilibrium constant $k(T,P)$, which is defined by the change in molar Gibbs potential $\Delta G^*(T,P)$ between the dissolved and the gaseous phases[1-4]. $R = 8.3144598\, Jmol^{-1}K^{-1}$ is the universal gas constant.

$$k(T,P) = exp\left(\frac{-\Delta G^*(T,P)}{RT}\right) \quad \text{(S3)}$$

The molar Gibbs potential is computed through the chemical potential $\mu(T,P)$ for both the dissolved (μ_{aq,w,O_2}) and gaseous (μ_{g,O_2}) O_2 phases.

$$G^*(T,P) = \left(\frac{\partial G}{\partial N}\right)_{T,P} = \mu(T,P) \quad \text{(S4)}$$

And hence $\Delta G^*(T,P) = \mu_{aq,w,O_2}(T,P) - \mu_{g,O_2}(T,P)$. Note that the chemical potential is a function of both temperature, T, and pressure, P. Hence, in order to compute the solubility of O_2 in water, we needed to know the chemical potential of O_2 in the gaseous and

dissolved aqueous phases as a function of temperature and pressure.

We referred to Equation S3, where the chemical potential is defined through the molar Gibbs potential, and to Equation S4 in order to compute $\mu(T,P)$ in relation to a reference state $\mu(T_{ref},P_{ref})$ - by calculating the differential $d\mu(T,P)$ from (T_{ref},P_{ref}) to (T,P) using the entropy S, the internal energy U, and the molar volume V for both phases.

$$d\mu(T,P) = dG^*(T,P) = d(U+PV-TS) = dU + VdP + PdV - TdS - SdT \quad \text{(S5)}$$

With the definition of the internal heat:

$$dU \equiv \delta Q - PdV \quad \text{(S6)}$$

and the definition for the entropy:

$$dS \equiv \frac{\delta Q}{T} \quad \text{(S7)}$$

we obtain:

$$d\mu(T,P) = \delta Q - PdV + VdP + PdV - TdS - SdT = VdP - SdT \quad \text{(S8)}$$

Therefore, we can rewrite the chemical potential as:

$$\mu(T,P) = \mu(T_{ref},P_{ref}) + \int_{P_{ref}}^{P} VdP' - \int_{T_{ref}}^{T} S(T')dT' \quad \text{(S9)}$$

The entropy S is a function of temperature and can be related to the specific heat at constant pressure $C_P(T)$, which is more useful (experimentally directly measurable) than the entropy itself, with Equation S7 and $\delta Q|_P = C_P dT$:

$$S(T) = S(T_{ref}) + \int_{T_{ref}}^{T} dS$$

$$= S(T_{ref}) + \int_{T_{ref}}^{T} \frac{C_P(T')}{T'} dT' + \overbrace{C_P(T) - C_P(T)}^{=0}$$

$$= S(T_{ref}) - C_P(T) + \overbrace{\frac{\partial}{\partial T}\left(T \int_{T_{ref}}^{T} \frac{C_P(T')}{T'} dT' \right)}^{\int_{T_{ref}}^{T} \frac{C_P(T')}{T'} dT' + C_P(T)} \quad \text{(S10)}$$

Therefore, we replaced in Equation S9 the integral over the entropy and obtained in Equation S11 the chemical potential at temperature T and reference pressure P_{ref}, which corresponds to the temperature-dependent function in Equation 2 from our Methods Section:

$$\tilde{\mu}(T) = \mu(T, P_{ref}) \quad \text{(S11)}$$

$$= \tilde{\mu}(T_{ref}) - S(T_{ref})[T - T_{ref}] + \int_{T_{ref}}^{T} C_P(T') dT' - T \int_{T_{ref}}^{T} \frac{C_P(T')}{T'} dT'$$

To fully solve for the chemical potential, we needed to take into account effects by pressure, described by $\int_{P_{ref}}^{P} V dP'$ in Equation S9: for the gas phase (hence for $\mu_{g,O_2}(T, P)$), we assumed an ideal gas with $PV = RT$ (for one mol) and hence:

$$\int_{P_{ref}}^{P} V dP' \Big|_g = RT ln\left(\frac{P}{P_{ref}}\right) \quad \text{(S12)}$$

For the dissolved phase, we assumed that the molar volume of dissolved O_2, V_{aq,O_2}, is smaller than that of water, V_w, and hence we get $V_{aq,O_2} = \lambda V_w$ with $0 < \lambda < 1$. Using this volume estimate with the standard molar volume for water at reference temperature and pressure (water is approximately incompressible for the pressures and temperatures that we

are interested in) of $V_w(T_{ref}, P_{ref}) = 18 \cdot 10^{-6} m^3 mol^{-1}$, we found that $\int_{P_{ref}}^{P} V_{aq,O_2} dP'$ is negligible and would, if included, only increase the solubility of O_2 by about 0.1%.

Therefore, from relationships S9, S11, and S12, we obtained Equation S13:

$$[\widetilde{O_2}]_{aq,w}(T,P) = \frac{f_{O_2} P}{P_{ref}} exp\left(\frac{\tilde{\mu}_{g,O_2}(T) - \tilde{\mu}_{aq,w,O_2}(T)}{RT}\right) \quad \text{(S13)}$$

which corresponds to Equation 1 in our Methods Section (compare also with other sources[1,3,4]). Note that the here derived pressure dependence corresponds to Henry's law.

From here, to compute $\tilde{\mu}_{g,O_2}(T) - \tilde{\mu}_{aq,w,O_2}(T)$, we needed standard values at (T_{ref}, P_{ref}) for the chemical potentials and entropy of gaseous and dissolved O_2 and the specific heat of O_2 in the gas phase, $C_{P,O_2,g}(T)$, and dissolved in pure water, $C_{P,O_2,aq}(T)$, as a function of temperature. Experiments[6] confirmed that $C_{P,O_2,g}(T) = C_{P,O_2,g}(T_{ref})$; all parameters at reference conditions can be found in Table 1S.

We examined the choice for the specific heat of dissolved O_2 in pure water in more detail below in the next Section 2.1 and the effects of salts on the solubility of O_2 in Sections 2.1.3 an 2.1.4.

2. Extended Discussion

2.1. Solubility

2.1.1. Specific heat of dissolved O_2 and solubility: tests and robustness of results

Experiments have shown $C_{P,O_2,aq}(T)$ to increase slightly as temperature decreases from 373-273 K, with no data available below this temperature range. There is reason to expect that the trend of increasing heat capacity continues from 273 K to ~225-235 K but possibly

also below this limit. Such a steep increase of $C_{P,O_2,aq}(T)$ is justified by arguing that the behavior of the specific heat of water at cooler temperatures below ~ 300 K directly corresponds to the behavior of the specific heat capacity of dissolved O_2. This assumption can be partially rationalized, as the smooth increase in the specific heat of pure water from 300 K into the supercooled water domain down to ~225-235 K can be explained by stronger hydrogen bonds at lower temperature[1], and hence a greater amount of heat being needed in order to increase the ambient temperature. This increase in hydrogen bonding for lower temperatures is expected to similarly affect the dissolved O_2, which will be partially polarized due to the water dipoles, leading potentially to a similar behavior of the specific heats of water and dissolved O_2 below 273 K. However, thermodynamics demands that $\lim_{T\to 0\ K} C_P = 0$. Indeed, some theoretical predictions find that the specific heat of supercooled water could start to decrease again for T~225-235 K, where potentially a liquid-liquid phase transition could occur[7-9]. We emphasize that experiments to date also allow for the heat capacity to continue increasing into much cooler regions below 235 K.

We used for our nominal best estimate (BE) the simplest assumption that $C_{P,O_2,aq}(T) = C_{P,O_2,aq}(T_{ref})$, but we also tested other reasonable forms of $C_{P,O_2,aq}(T)$ (all show a similar result), and derived a thermodynamic worst case, which technically cannot be reached – assuming the validity of Equations 1 and 2 from our Methods Section.

Assuming $\left(\frac{C_{P,O_2,aq}}{C_{P,w}} \sim const.\right)$ below 273 K (specific heat of dissolved O_2 scales with the specific heat of water) and taking the predicted specific heat behavior[7-9] of supercooled water – consisting of an initial power law increase below 273 K towards a critical temperature, T_{crit}, of 225-235 K below which the heat capacity exponentially decays

towards 0 K, we found that at 140 K O_2 solubility values are between 1.2 times smaller or larger than our best estimate (BE). Therefore, if the specific heat of dissolved O_2 behaves similar to the theoretical predictions for the specific heat of supercooled water, then our best estimate is an excellent prediction for the solubility of O_2 below 273 K.

Alternatively, we examined what happens if we had assumed that instead the specific heat of dissolved O_2 in water behaves similar to the specific heat of a "normal fluid" that does not show anomalous behavior like water, such as a very salty NaCl brine[9]. For NaCl-brines[9] and many other brines including perchlorates[10], the specific heat is rather constant but does decay slowly with decreasing temperature. Assuming an unusually strong linear decrease of $C_{P,O_2,aq}(T)$ by 50% from 293-140 K (a few factors to ~1 order of magnitude larger than generally found for very salty brines[9,10]), we obtained values for the solubility of O_2 at 140 K that were ~3 times smaller than our best estimate. Thus, we concluded that, for all reasonable forms of $C_{P,O_2,aq}(T)$, the solubility values at 140 K are generally similar to the best estimate assuming a constant heat capacity for dissolved O_2 in pure water. In the next section, we derived a thermodynamic lower bound that cannot be reached.

2.1.2. Specific heat of dissolved O_2 and thermodynamic worst case (WC)

The solubility of O_2 in mol m^{-3} in pure water or brine depends on the specific heat of dissolved O_2, $C_{P,O_2,aq}$, in the following way (see Equations 1 and 3 in Methods Section, and Equations S11 and S13):

$$[O_2]_{aq}\left(T,\zeta(T)\right) = F(T)exp\left(\frac{\zeta(T)}{RT}\right)$$

with $\zeta(T) = -\int_{T_{ref}}^{T} C_P(T')dT' + T\int_{T_{ref}}^{T} \frac{C_P(T')}{T'}dT'$ (S14)

The function $\zeta(T)$ contains all the uncertainties introduced by the behavior of $C_{P,O_2,aq}(T)$ below 273 K, whereas $F(T)$ depends only on well-constrained properties and monotonically increases for smaller temperatures. We developed a thermodynamic lower limit for $\zeta(T)$, which gives a lower limit for the solubility of O_2 in supercooled water. To achieve this, we needed to assess the behavior of $\zeta(T)$.

First, we showed that $\zeta(T_1) > \zeta(T_2)$: $\forall(T_1 < T_2)$, so independent of the temperature dependence of the heat capacity, solubilities monotonically increase for smaller temperatures, with the minimum solubility curve for $T < 273\ K$ given by:

$[O_2]_{aq}\big(T, \zeta(T)\big) > [O_2]_{aq}(T)\big|_{min} = [O_2]_{aq}(T, \zeta^*)$ with $\zeta^*(T) = \zeta(T = 273\ K)$ (S15)

Therefore, the integral in the exponent, which is monotonically increasing for smaller T is cut at the lowermost temperature where $C_{P,O_2,aq}(T)$ is exactly known, at 273 K. Equation S15 was derived in the following: $\forall T: T > 0, C_P(T) > 0$, therefore, because of $T' > T$ and therefore $\frac{T}{T'} < 1$, we have:

$$\left\| T\int_{T_{ref}}^{T} \frac{C_P(T')}{T'}dT' \right\| < \left\| \int_{T_{ref}}^{T} C_P(T')dT' \right\| \quad \text{(S16)}$$

So, for $T_{ref} > T$, we concluded that:

$$\int_{T_{ref}}^{T} C_P(T')dT' < 0 \text{ and } \int_{T_{ref}}^{T} \frac{C_P(T')}{T'}dT' < 0 \quad \text{(S17)}$$

And hence $\zeta(T) = -\int_{T_{ref}}^{T} C_P(T')dT' + T\int_{T_{ref}}^{T} \frac{C_P(T')}{T'}dT' > 0, \forall\big(T, C_P(T)\big)$ (S18)

Next to $\zeta(T)$ being a positive function, as shown in Equation S18, $\zeta(T)$ also monotonically increases for smaller *T* because:

$$\zeta(T_1) = \zeta(T_2) \overbrace{- \int_{T_2}^{T_1} C_P(T')dT' + T \int_{T_2}^{T_1} \frac{C_P(T')}{T'} dT'}^{">0" \; as \; shown \; in \; S18} > \zeta(T_2), \forall (T_1 < T_2) \quad \text{(S19)}$$

With Equation S19, we have proven the lower bound for the solubility of oxygen shown in Equation S15. We showed in Figure 1S the worst case in comparison to the best estimate shown in Figure 1 for pure supercooled water.

Note that, formally, the worst case cannot be reached as it demands that $C_P(T) = 0$, and this is only possible for $T = 0\ K$. Hence, our worst case is indeed an unapproachable limit; however, it provides the logic for a lowermost bound on O_2 solubility, and it is important to note that the true solution is likely to be much greater.

2.1.3. Derivation of Pitzer coefficients for perchlorates

We used the experimental results[11-13] on the O_2 solubility in perchlorate brines containing the salts $NaClO_4$, $KClO_4$, $RbClO_4$, and $LiClO_4$ to derive the Pitzer interaction coefficients for the O_2-perchlorate ion interaction O_2-ClO_4, described by $\lambda_{O_2-ClO_4^-}$ in Equation 4 in our Methods Section. Generally, the temperature dependence of the Pitzer coefficients is negligible, but we examined this in the next section. The O_2-cation interaction coefficients, λ_{O_2-c}, for the cations c=(Ca^{2+}, Mg^{2+}) were taken from literature[11].

We showed the results for the salting out coefficient $\gamma_{O_2}(X, m_X)$ for a brine containing m_X mol kg^{-1} of the salt X in Figure 2S for chlorides and perchlorates (compare with Equation 3 in our Methods Section). There were no direct data on O_2 solubility for Ca- or Mg-perchlorates, and hence there are no data on the secondary Pitzer interaction

coefficients describing the interactions between O_2-cation-ClO_4^-, λ_{O_2-c-a}. However, as we explained in the Methods Section, the secondary interaction coefficients are generally negative as they represent the disturbance of the respective cation-O_2 and anion-O_2 fields, which drive the solubility of O_2 in the liquid. Hence any reasonable value for $\lambda_{O_2-c-a} < 0$ would only lead to a reduction of the salting out factor $\gamma_{O_2}(X, m_X)$, and would hence result in *greater* solubilities for O_2 in perchlorate brines. Thus, we assume $\lambda_{O_2-c-a} = 0$ for Ca- and Mg-perchlorate brines—and accept that this approach will tend to, if anything, underestimate the solubilities for O_2 in those brines.

2.1.4. Temperature dependence of Pitzer coefficients

It is commonly assumed that the Pitzer coefficients introduced in Equation 4 in the Methods Section are not temperature-dependent[14,15]. Nonetheless, the wide amount of experimental data that we collected and tested our results against[11-13,15-19] suggested a slight increase of Pitzer coefficients for smaller temperatures, and thus we labored to find a way to estimate how much this might change the estimate for the O_2 solubilities in Ca- and Mg-perchlorate brines by neglecting a potential temperature dependence of the Pitzer coefficients. We found that the temperature dependence of the Pitzer coefficients could lead to an additional decrease of the solubility of O_2 in Ca- and Mg-perchlorate brines at 140 K by a factor of 3-10 (most likely less than 5).

To derive this estimate, we made use of the salting out theory of Tromans[3]: he observed that a salt reduced the solubility of O_2 by reducing the molar volume of water (that is dissolving O_2) into an apparent water volume V_{app}. He also found that the change in

apparent water volume could be used to infer the salting out factor. Thus, what we did first was to derive the apparent molar water volume as a function of the brine density ρ_{brine}, molar concentration m_X of the salt X with molecular mass M_X in kg mol^{-1} and molecular crystalline volume V_X, and the molar mass of pure water M_w - leading to Equation S20:

$$V_{app}(T) = \left(\frac{1 + m_X M_X}{\rho_{brine}(T)} - m_X V_X\right) M_w(T) \quad \text{(S20)}$$

Equation S20 was easiest to derive by computing the density of the brine ρ_{brine} assuming that water has an apparent volume V_{app} and that the salt preserves its crystalline molecular volume V_X, and solving for V_{app} (for values, see Table 2S).

Second, we computed how this apparent water volume in Equation S20 changes during a reduction of temperature from 298-140 K. The temperature dependence for $V_{app}(T)$ results from the temperature dependence of the brine density and the molar volume of water, but $V_{app}(T)$ is mainly sensitive to small changes in brine density $\rho_{brine}(T)$. Thus, to compute how $V_{app}(T)$ changes with temperature, we focused solely on determining how the density of a brine, $\rho_{brine}(T)$, changes with temperature when salt and molality are fixed.

There is no data or theoretical predictions for the temperature dependence of the density of perchlorate brines, nor is data available on their densities under ambient conditions. To obtain an estimate of how much the density changes for Ca- and Mg-perchlorate brines with up to ~4.2 mol of salt per kg water (= 4.2 m, the concentration at the eutectic for Ca-perchlorate[10,14,20,21], see Table 3S for values) for a temperature change

from 298-140 K, we turned towards experimental data[22,23] on the density of a range of brines of different compositions (i.e., $NaCl, MgCl_2, CaCl_2, KCL, K_2SO_4, MgSO_4,$ and Na_2SO_4), and examined how their densities changed from 373-243 K for various salt concentrations up to 30 weight % (or ~4.2 m). We observed that variations in brine density driven by temperature are rather small and would amount to less than a 5-10% increase in density from 298-140 K—assuming a linear increase of the density towards smaller temperatures with a gradient determined between 303-243 K. Note also that our available data indicates that the density variations decrease for smaller temperatures, and one obtains very similar results including data up to 373 K in order to estimate the density gradient with temperature.

The derivation of the density gradient is illustrated in Figure 3S a, where we plotted the density gradient for diverse brines with temperature as a function of molality using density data from 303-243 K. Extrapolating these values would lead to a density increase by less than 5-10% at 140 K in comparison to 298K (on average about 6% for a random sampling of temperature intervals between 303-243 K). Note that this compares excellently with the predicted and measured density variation of NaCl brines between 240-300 K[9], which would lead to a 5% increase in the NaCl brine density from 298-140 K for a linear extrapolation (which most likely overshoots the density change).

With this approach, we used a density increase of 5-10% for a $Ca(ClO_4)_2$ brine from 298-140 K to measure the apparent volume changes for a 4.2 m $Ca(ClO_4)_2$ brine from 298-140 K, $\frac{V_{app}(\varepsilon \cdot \rho_{brine})}{V_{app}(\rho_{brine})} \approx \frac{V_{app}(140\ K)}{V_{app}(298\ K)}$, where ε is either 1.05 or 1.1 (corresponding to the increase in brine density from 298-140 K by 5-10%). This is illustrated in Figure 3S b,

where we plotted the change in apparent water volume as a function of standard brine density for our 4.2 m $Ca(ClO_4)_2$ brine at 298 K. This density was calculated using Equation S20, inverting for the brine density and setting $V_{app}(298\ K) \approx V_w(298\ K)$ as a first order approximation. This leads to an estimated 4 m $Ca(ClO_4)_2$ brine density at 298 of ~1440 kgm^{-3}, and thus values for $\frac{V_{app}(\varepsilon \cdot \rho_{brine})}{V_{app}(\rho_{brine})}$ between 0.875-0.935.

We then used the theory of Tromans[3] that can be used to infer how a decrease in apparent volume by a factor of 0.875-0.935 affects the increase in salting out coefficient. Adhering to the reference KOH line of Tromans[3] (see his Figure 7 which shows how the salting out factor is a function of apparent volume change), we saw that a temperature change from 298-140 K would maximally yield an increase in salting out factor between ~3-10 (for the average value of 6% density increase from 298-140 K, it would be a factor of ~5). Therefore, even if we accounted for a potential temperature dependence of Pitzer coefficients, then we would maximally obtain dissolved O_2 concentrations at 140 K that are 3-10 (and likely less than ~ 5) times smaller than what we observed with our best estimate shown in Figures 1-3.

2.1.5. Stability of perchlorate brines under Martian pressures

For a brine at surface temperature, T, and pressure, P, to be liquid, the concentration of salt must correspond to the critical molality $m_X(T)$ at this temperature (see methods) but the pressure must be also above the triple point pressure $P_{TP}(m_X)$ for that specific brine. For pure water, the triple point pressure is ~6.1 mbar, around the average atmospheric pressure on Mars. At this pressure, the melting and boiling temperatures converge, which would

limit the presence for liquid water to temperatures very close to 273 K (this explains the narrow water range in Figure 1c, d).

For brines, however, the triple point pressure is reduced. Specifically, for Ca-perchlorate brines, the salt concentrations m_X needed for the brine to remain liquid at the surface temperatures obtained with our GCM simulations (below ~230 K, see Table 4S) result in triple point pressures at least one order of magnitude below the minimal atmospheric surface pressure of ~1 mbar encountered in our GCM. The triple point pressure can be approximately computed as the intersection between the standard "water ice-vapor" phase curve in temperature-pressure space and an isotherm for a given temperature $T_{TP} = max(T, T_{eu})$ (with surface temperatures T ranging from ~145-230 K and the eutectic temperature for Ca-perchlorate brines of $T_{eu} = 198.2\ K$). Therefore, the pressure conditions on Mars support the studied Ca-perchlorate brines with salt concentrations $m_X(T)$, which we show in Figure 2, to be liquid. Moreover, due to the approximate incompressibility of water and brines between ~1 bar and ~1 mbar, we do not expect any significant change in eutectic temperature from Earth to Mars surface pressures.

2.2. Climate model

2.2.1. Albedo variation with obliquity change

The albedo of the Martian surface and how it changes with obliquity can potentially be a significant control on surface temperatures. For any rotating planet, annually averaged insolation at the poles increases going from zero obliquity to larger values, leading to a warming of cooler regions (poles) and a cooling of warmer regions (low latitudes) as

obliquity rises (as seen in Table 4S for annually averaged temperatures for present-day Mars). For a homogenous obliquity across the planet's surface, above 54° obliquity, the poles receive more insolation than the equator,[24] and become the 'warmer' regions of the planet on an annual basis, whereas the tropics become the 'cooler' region and, therefore, the preferred location for surface ice formation.

The exact timing of how surface ice on Mars evolves with changing obliquity is uncertain, as is the exact distribution of tropical surface ice at high obliquity. Different studies have alternatively suggested that water ice will condense widely across high topographic and/or high thermal inertia sites on the planet at high obliquity[25], or only in localized glacial deposits on the flanks of the Tharsis Montes[26]. The impact of the exact distribution on surface albedo on temperature can be not insignificant. For example, during the early phases of an obliquity transition (say, from lower to higher values), ice will likely remain in the polar regions for some time after the obliquity rises because of the multi-kilometer thickness of the polar caps. Albedo, then, may remain relatively high at the poles, while it concurrently increases in the tropics due to the development of ice deposits there as obliquity rises. Later in the obliquity transition, though, it is likely that the albedo at the poles will decrease as a darker, dirty sublimation lag forms atop the retreating polar ice cap. The timing of this transition will be gradual and difficult to model precisely.

We performed a series of tests of the end-member cases from previous studies[25,26], looking at the effect of their putative ice distributions on the global surface temperatures at high obliquity, looking at different stages of the aforementioned evolution, with bright ice in the poles only, in both the tropics and poles, and in the tropics only. While the annually

averaged global surface temperature does decrease with the presence of tropical ice, (which covers a larger fraction of the surface with high albedo ice than the poles at low obliquity), the general trend in surface temperatures with obliquity continues regardless of the exact distribution of ice, and global mean temperatures continue decreasing with rising obliquity. Differences in surface temperatures between the end members of surface ice distribution are in the range of 5-10 K. This behavior is reflected in the O_2 solubility results in Figure 3a, where, with increasing obliquity, the maximum solubility (corresponding to the lowest annually averaged temperature on the surface) continuously increases.

2.2.2. South pole ice caps

We did not impose any constraints on ice caps for our obliquity calculations; however, on modern-day Mars, observations reveal a perennial polar cap of CO_2 ice in the south, which global climate models cannot self-consistently reproduce without making specific, ad hoc assumptions. In our GCM calculations, we set the surface temperature to the pressure-dependent CO_2 frost point for all locations poleward of -85°. This GCM model is listed in Table 4S as "25° with SPC" and represents the most realistic climate model for modern-day Mars. This model has been used to study modern-day Mars and to create Figure 2 in our main manuscript. Mechanisms for forming such a south polar CO_2 cap are not fully understood, and hence we cannot (and do not) extrapolate it to other obliquities; therefore, we also run a 25° obliquity simulation without the assumption of a south polar cap in order to be able to better grasp how O_2 solubilities change with obliquity, as shown in Figure 3. Note that Figure 3 highlights that the "with SPC" model leads to about one order of

magnitude greater maximum solubility values on the Martian surface today because of the fixed low temperatures for the southern polar regions.

2.2.3. Atmospheric collapse

For obliquities below ~10°, we found that, on geological timescales, the atmosphere collapses due to the presence of permanent CO_2 cold traps in the polar regions. This can be seen in Table 4S, where we show how the minimum value for the annually averaged surface temperature is pegged to the CO_2 frost point temperature for smaller obliquities. This model scenario is not likely to have occurred in the last twenty million years, might occur rarely in the next five million years (see Figure 3 b), but might have occurred more frequently in the deeper past[27].

Atmospheric collapse will, hence, not significantly affect our results, but it is interesting to consider what could happen to O_2 concentrations during such an event. O_2 is produced in the Martian atmosphere through photochemical reactions from CO_2 and H_2O, and subsequent hydrogen escape[28]; hence, in order to understand how O_2 reacts to atmospheric collapse, we needed to understand how atmospheric water and CO_2 behave during an atmospheric collapse. Atmospheric water will freeze out at the poles due to their lower surface temperatures at smaller obliquities. This will increase O_2 levels[29]; however, for a reduction of water levels by 1-4 orders of magnitude O_2 levels are only increased by a factor of about three[29] – so this effect will be minor. The dominant effect will be the removal of atmospheric CO_2 and, hence, the major source for O_2 formation. It is therefore likely that during an atmospheric collapse O_2 levels will significantly drop.

3. Validity of averaging method and extension to daily temperature variations

For practical reasons, we generated climate data of annual averages of temperature and pressure as a function of location on the Martian surface for varied obliquities (see details on climate model above). The solubility at each point on Mars' surface was evaluated for this annual average of temperature and pressure. It is important to point out that by doing such a time-averaging, the solubility results shown in Figures 1-3 and the differences between minimum and maximum solubilities shown in Figure 3 are smaller than if we instead would have computed the annual average of solubility as a function of hourly or daily averages.

With our averaging approach, we underestimated the solubility because $[O_2]_{aq}$ increases exponentially towards 0 K, hence the gradient $\left|\frac{\partial [O_2]_{aq}}{\partial T}\right|$ is greater for smaller temperatures and the solubility evaluated at the average annual temperature is smaller than the solubility averaged over a greater time (and hence temperature) interval. The solubility is linear in pressure and thus using the annually averaged pressure does represent the precise average for the solubility as a function of pressure. Computing the annual average of the O_2 solubility directly from the non-averaged temperature (daily or hourly) would only increase our O_2 solubility estimates and strengthen the conclusions.

4. Spatial gradients, oxidative weathering, and next steps

As we showed in Figure 4Sa, it is primarily temperature and secondarily pressure that determines the potential solubility of O_2 in Martian surface and near surface environments.

The geographic variations in O_2 content lead to natural gradients in concentrations across environments where waters might naturally flow[30,31]. Areas where spatial gradients (as shown in Figure 4s b) in O_2 solubility are largest are observed on Tharsis, across Meridiani, Arabia Terra, and within the Hellas basin.

The opportunity for oxidative processes involving O_2 during the chemical weathering of Mars' crust will reflect a convolution between the availability of water and its O_2 content. Our results suggest that we should expect a degree of patchiness in oxidative chemistry associated with weathering; the same would be true for the potential energetic gradients for aerobic respiration.

We have computed the climatically-induced potential for aerobic environments on Mars in 3D (Figure 2), assuming that perchlorate and water are equally available everywhere close to the planet's surface. Naturally, we expect spatial variability in the distribution of water and perchlorates, and hence as a next step, it will be inevitable to explore how perchlorate and water distribution across the Martian surface might vary and how this will affect our results. This, however, demands a much deeper insight into the dynamics of brine formation/destruction and gas dissolution processes, which are far beyond the scope of this first, critical piece of work. It would be also interesting to take into account global Mars Odyssey GRS data on chlorine and hydrogen – however data are only available down to a few microns in depth, are strongly modulated by dust, and cannot see the meaningful depths of the regolith that thermocycle daily or seasonally.

Figures SOM

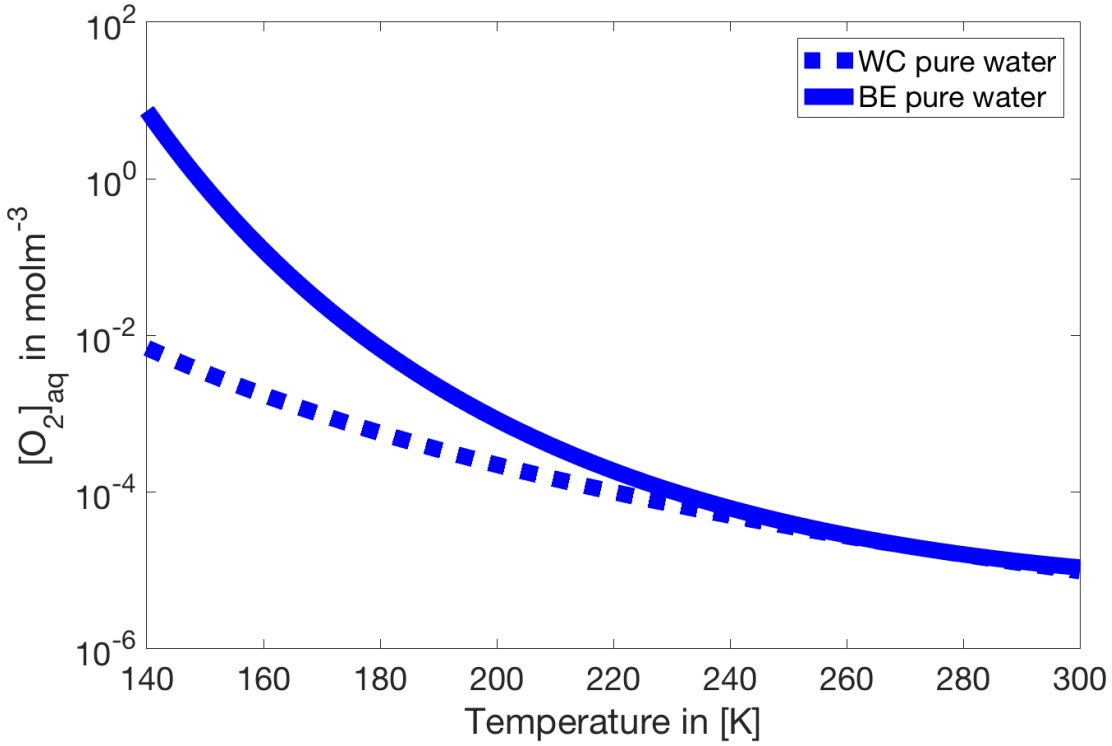


Figure 1S – Thermodynamic lower limit to the solubility of O_2: we compared the best estimate (BE) for the solubility curve of O_2 in supercooled water to the thermodynamic worst case (WC), which sets the specific heat of dissolved O_2 to zero for temperatures below 273 K. Formally, this is only possible for T = 0 K, and thus WC marks a lowermost estimate.

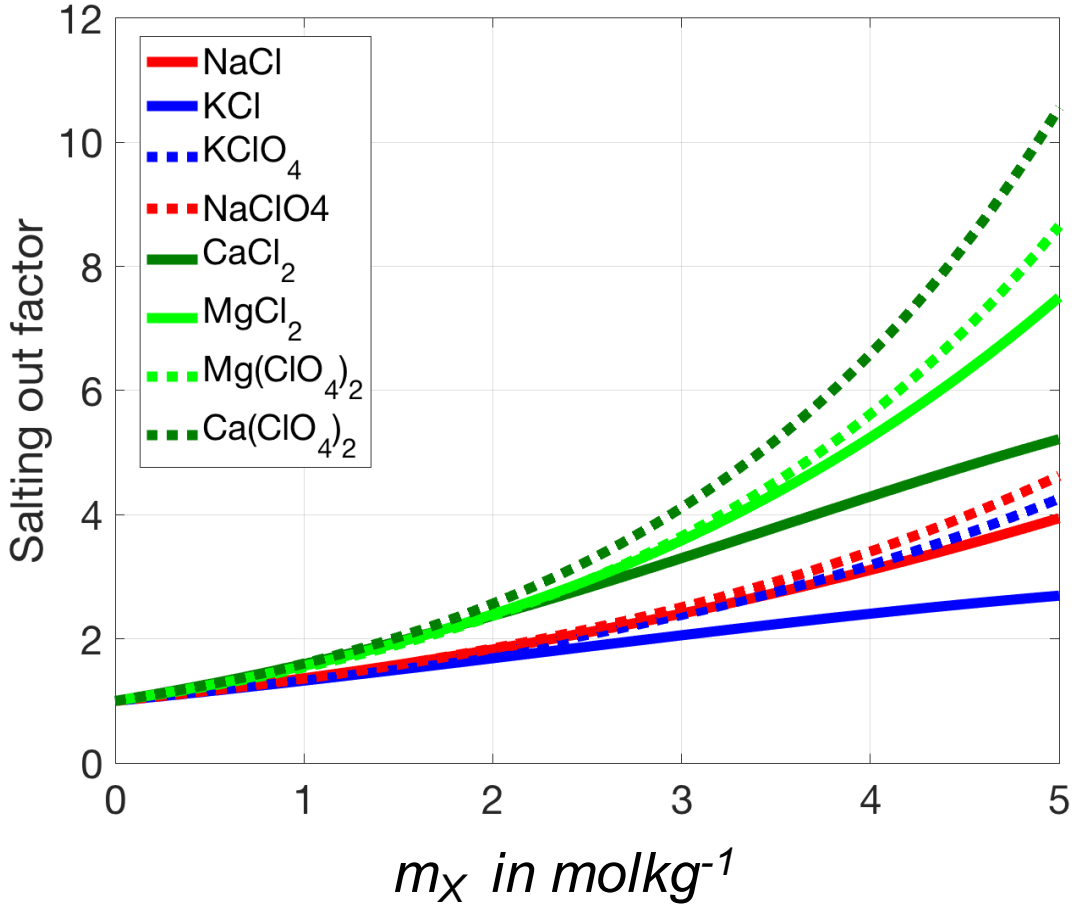


Figure 2S – Salting out coefficients for O_2 in perchlorate and chloride brines: using data on the solubility of O_2 in various chlorides (solid) and perchlorate (dashed) brines[11-13], we computed Pitzer interaction coefficients at 298 K for Ca- and Mg-perchlorates, describing the salting out factor $\gamma_{O_2}(X, m_X)$ of O_2 as a function of salt concentration. We also plotted the salting out factors for other relevant brines (for parameters, see Table 2S). $[O_2]_{aq,X}(T,P) = \frac{1}{\gamma_{O_2}(X,m_X)}\,[O_2]_{aq,w}(T,P)$, wherein the salting out factor relates the solubility of O_2 in pure water $[O_2]_{aq,w}(T,P)$ and in the brine $[O_2]_{aq,X}(T,P)$. The salting

out factors for Mg- and Ca-perchlorates are likely overestimated here, as we neglected electrostatic cation-anion-O_2 interactions.

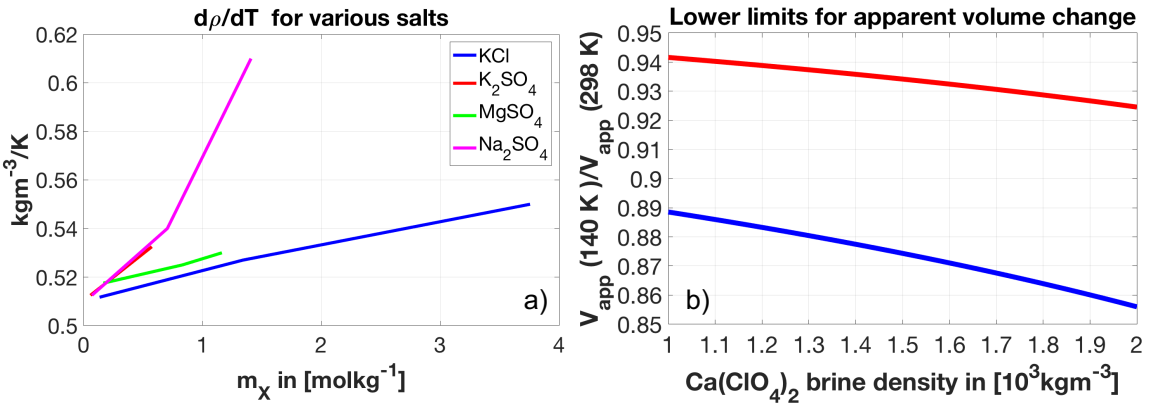


Figure 3S – The temperature dependence of salting out factors for O_2 in various brines: a) the density gradient with temperature as a function of molality using density data from 303-243 K. **b)** Using an average density variation from 5-10% for a $Ca(ClO_4)_2$ brine, we computed how the apparent volume of water changes from 298-140 K for a range of 4.2 m $Ca(ClO_4)_2$ brine densities at 298 K (our estimated value is ~1400 kgm^{-3}, see below) by plotting $\frac{V_{app}(\varepsilon \cdot \rho_{brine})}{V_{app}(\rho_{brine})} \approx \frac{V_{app}(140\ K)}{V_{app}(298\ K)}$, where ε is either 1.05 (red) or 1.1(blue) (representing the increase in brine density from 298-140 K by 5-10%). The apparent volume depends slightly on the reference brine density at 298 K, which we estimated by using Equation S20, inverting for the brine density and setting approximately $V_{app}(298\ K) \approx V_w(298\ K)$, ultimately leading to an estimated 4.2 m $Ca(ClO_4)_2$ brine density at 298 K of ~1440 kgm^{-3}, and $\frac{V_{app}(\varepsilon \cdot \rho_{brine})}{V_{app}(\rho_{brine})}$ between 0.875 and 0.935. Along the KOH line in the Tromans model[3], we get an increase of the salting out factor between ~3-10 (and likely less than ~ 5).

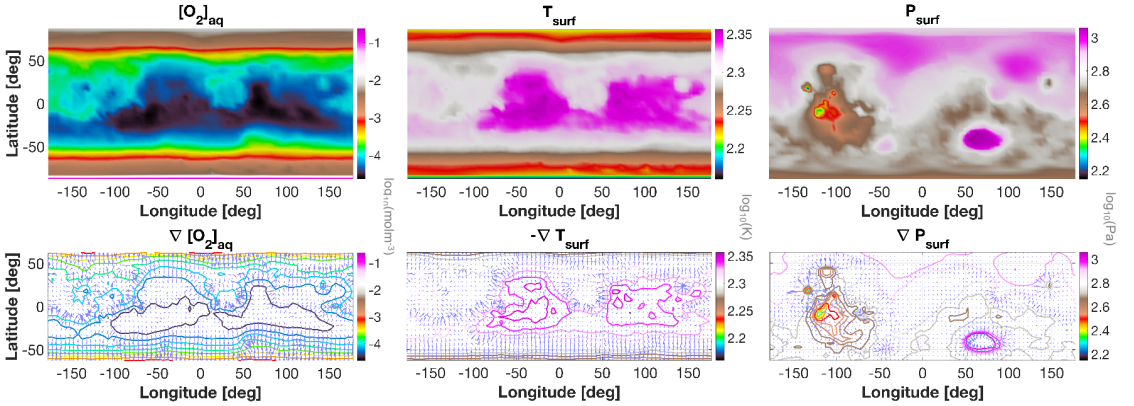


Figure 4S – The predominant factors that control O_2 solubility, and spatial O_2 solubility gradients on modern-day Mars. This plot shows the O_2 solubilities for modern-day Mars using local annual averages for surface temperature and pressure **(top)** as well as spatial gradients **(bottom)** in O_2 solubility. The primary control on O_2 solubility is temperature and secondary modifications result from pressure.

TABLES

Thermodynamic parameters for the solubility of O_2 in pure water					
Phase of O_2	$\tilde{\mu}(T_{ref})$ *[J mol^{-1}]*	$S(T_{ref})$ *[J K^{-1} mol^{-1}]*	$C_P(T_{ref})$ *[J K^{-1} mol^{-1}]*	T_{ref} *[K]*	P_{ref} *[Pa]*
Gaseous	0	205.028	29.332	298	$1.01325 \cdot 10^5$
Aqueous	16506	109	205.266	298	$1.01325 \cdot 10^5$

Table 1S: Lists all thermodynamic parameters needed to compute the specific heat capacity for O_2 at constant pressure in the gaseous and aqueous phases, essential for deriving the solubility of O_2 in pure water as a function of pressure and temperature in Equations 1 and 2. The parameters are taken from experiments[1]. The partial volume fraction in the Martian atmosphere is approximately $f_{O_2} = 0.00146$[5], and $R = 8.3144598\ \mathrm{Jmol^{-1}K^{-1}}$ is the universal gas constant.

Salting-out parameters								
Salt	**Ions**	f_c	f_a	λ_{O_2-c} *[kg mol^{-1}]*	λ_{O_2-a} *[kg mol^{-1}]*	λ_{O_2-c-a} *[kg^2 mol^{-2}]*	M_X *[kg mol^{-1}]*	V_X *[10^{-6} m^3 mol^{-1}]*
$Ca(ClO_4)_2$	Ca^{2+}/ClO_4^-	1	2	0.2497	-0.007	0	0.23898	90.147
$Mg(ClO_4)_2$	Mg^{2+}/ClO_4^-	1	2	0.2298	-0.007	0	0.22321	101
$NaClO_4$	Na^+/ClO_4^-	1	1	0.1602	-0.007	0	0.12244	48.988
$KClO_4$	K^+/ClO_4^-	1	1	0.1519	-0.007	0	0.13855	54.980
$MgCl_2$	Mg^{2+}/Cl^-	1	2	0.2298	0	-0.00565	0.09521	40.81
$CaCl_2$	Ca^{2+}/Cl^-	1	2	0.2497	0	-0.0169	0.11098	50.5
NaCl	Na^+/Cl^-	1	1	0.1602	0	-0.00919	0.05844	27.02
KCl	K^+/Cl^-	1	1	0.1519	0	-0.0211	0.07455	37.52
$MgSO_4$	Mg^{2+}/SO_4^{2-}	1	1	0.2298	0.0878	0	0.12037	40.7
K_2SO_4	K^+/SO_4^{2-}	2	1	0.11519	0.0878	0	0.17426	65.48
Na_2SO_4	Na^+/SO_4^{2-}	2	1	0.1602	0.0878	-0.046	0.14204	53.33

Table 2S: Specifies the ions, molality of cations (c) or anions (a) per molality of salt, the Pitzer coefficients λ[11-13] for the interaction of O_2 molecules with cations (c), anions (a), and cations & anions (c-a), and the crystalline molar mass M_X and volume V_X (values have been taken from experiments[3] or computed using data[22] with $V_X = M_X/\rho_X$, where ρ_X is the density at 298 K and 1 atm~1.01325 bar) for $Ca(ClO_4)_2$, $Mg(ClO_4)_2$, $NaClO_4$, $KClO_4$,

$MgCl_2$, $CaCl_2$, NaCl, KCl, $MgSO_4$, K_2SO_4, and Na_2SO_4. The crystal data is only necessary to estimate the temperature dependence of the salting out factor in the SOM.

Melting curve parameters						
Salt	T_{eu} *[K]*	m_{eu} *[mol kg^{-1}]*	p_3 *[10^{-5}]*	p_2	p_1	p_0
$Ca(ClO_4)_2$	198.2	4.176	-1.0689	0.0069556	-1.5378	119
$Mg(ClO_4)_2$	209.3	3.375	-1.4134	0.0094903	-2.1498	167.22
$NaClO_4$	239.2	9.2	-0.6053	0.003012	-0.6241	69.098
$MgCl_2$	240.15	2.84	-8.5302	0.063718	-15.911	1330.6
$CaCl_2$	224	4	-2.56	0.017736	-4.1378	328.69
NaCl	251.85	5.17	0	0	-0.23877	65.22
$MgSO_4$	269.55	1.72	0	0	-0.47778	130.51

Table 3S: Specifies the eutectic temperature T_{eu}, the eutectic salt concentration m_{eu} (in mol salt per kg water), and the parameters used to parameterize the critical concentration for the melting curve with temperature for the salts shown in Figure 1, which are $Ca(ClO_4)_2$, $Mg(ClO_4)_2$, $NaClO_4$, $MgCl_2$, $CaCl_2$, NaCl, and $MgSO_4$. Values are taken from experimental data[10,32]. The melting curve was defined by $m_X(T) = \sum_{i=0}^{3} p_i T^i \ , T > T_{eu}$ and $m_X(T) = m_{eu} \ , T \leq T_{eu}$, and determined with experimental data[10,14,20,21,32,33].

Temperature variation on the Martian surface with obliquity			
Obliquity *[°]*	**Minimum annual average** *[K]*	**Global annual average** *[K]*	**Maximum annual average** *[K]*
5	144.4	205.7	229.7
10	144.5	205.4	229.5
15	149.6	205.0	229.0
20	156.0	204.4	228.3
25	161.8	203.703	227.413
25 (with SPC)	144.2	203.651	227.389
40	172.1	201.6	224.1
60	179.7	198.0	214.4
90	177.2	195.4	208.8

Table 4S – Surface temperatures across obliquities: for the simulated obliquities, the minimum, average, and maximum values on the Martian globe for annually averaged surface temperatures. The annotation "with SPC" is representing modern-day Mars, where a perennial CO_2 cap at the south pole was assumed.

Extended References

1 Tromans, D. Temperature and pressure dependent solubility of oxygen in water: a thermodynamic analysis. *Hydrometallurgy* **48**, 327-342 (1998).

2 Tromans, D. Oxygen solubility modeling in inorganic solutions: concentration, temperature and pressure effects. *Hydrometallurgy* **50**, 279–296 (1998).

3 Tromans, D. Modeling Oxygen Solubility in Water and Electrolyte Solutions. *Ind. Eng. Chem. Res.* **39**, 805-812 (2000).

4 Taylor, D. F. Thermodynamic Properties of Metal-Water Systems at Elevated Temperatures. *Journal of The Electrochemical Society* **125**, 808, doi:10.1149/1.2131553 (1978).

5 Mahaffy, P. R. *et al*. Abundance and isotopic composition of gases in the martian atmosphere from the Curiosity rover. *Science* **341**, 263-266, doi:10.1126/science.1237966 (2013).

6 Chase, M. W. J. *et al*. in *Standard Reference Data Program* Vol. 1.0 (National Institute of Standards and Technology, Gaithersburg, MD 20899, 1986).

7 Holten, V., Palmer, J. C., Poole, P. H., Debenedetti, P. G. & Anisimov, M. A. Two-state thermodynamics of the ST2 model for supercooled water. *J Chem Phys* **140**, 104502, doi:10.1063/1.4867287 (2014).

8 Holten, V., Bertrand, C. E., Anisimov, M. A. & Sengers, J. V. Thermodynamics of supercooled water. *J Chem Phys* **136**, 094507, doi:10.1063/1.3690497 (2012).

9 Archer, D. G. & Carter, R. W. Thermodynamic Properties of the NaCl + H2O System. 4. Heat Capacities of H2O and NaCl(aq) in Cold-Stable and Supercooled States. *J. Phys. Chem. B* **104**, 8563-8584 (2000).

10 Toner, J. D. & Catling, D. C. Water activities of NaClO4, Ca(ClO4)2, and Mg(ClO4)2 brines from experimental heat capacities: Water activity >0.6 below 200K. *Geochimica et Cosmochimica Acta* **181**, 164-174, doi:10.1016/j.gca.2016.03.005 (2016).

11 Clegg, S. L. & Brimblecombe, P. The solubility and activity coefficient of oxygen in salt solutions and brines. *Geochimica et Cosmochimica Acta* **54**, 3315-3328 (1990).

12 Konnik, E. I. Salting-out and Salting-in of Gaseous Non-electrolytes in Aqueous Solutions of Electrolytes. *Russian Chemical reviews* **46**, 577-588 (1977).

13 Khomutov N. E., K. E. I. Solubility of oxygen in aqueous electrolyte solutions. *Russian Journal of Physical Chemistry* **48** (1974).

14 Toner, J. D., Catling, D. C. & Light, B. A revised Pitzer model for low-temperature soluble salt assemblages at the Phoenix site, Mars. *Geochimica et Cosmochimica Acta* **166**, 327-343, doi:10.1016/j.gca.2015.06.011 (2015).

15 Silvester, L. F. & Pitzer, K. S. Thermodynamics of electrolytes. X. Enthalpy and the effect of temperature on the activity coefficients. *Journal of Solution Chemistry* **7**, 327–337, doi:10.1007/BF00662893 (1978).

16 Millero, F. J., Huang, F. & Laferiere, A. L. The solubility of oxygen in the major sea salts and their mixtures at 25°C. *Geochimica et Cosmochimica Acta* **66**, 2349-2359, doi:10.1016/s0016-7037(02)00838-4 (2002).

17 Millero, F. J. & Huang, F. Solubility of Oxygen in Aqueous Solutions of KCl, K2SO4, and CaCl2as a Function of Concentration and Temperature. *Journal of Chemical & Engineering Data* **48**, 1050-1054, doi:10.1021/je034031w (2003).

18 Geng, M. & Duan, Z. Prediction of oxygen solubility in pure water and brines up to high temperatures and pressures. *Geochimica et Cosmochimica Acta* **74**, 5631-5640, doi:10.1016/j.gca.2010.06.034 (2010).

19 Weiss, R. F. The solubility of nitrogen, oxygen and argon in water and seawater. *Deep-Sea Research* **17**, 721-735 (1970).

20 Toner, J. D., Catling, D. C. & Light, B. The formation of supercooled brines, viscous liquids, and low-temperature perchlorate glasses in aqueous solutions relevant to Mars. *Icarus* **233**, 36-47, doi:10.1016/j.icarus.2014.01.018 (2014).

21 Toner, J. D., Catling, D. C. & Light, B. Modeling salt precipitation from brines on Mars: Evaporation versus freezing origin for soil salts. *Icarus* **250**, 451-461, doi:10.1016/j.icarus.2014.12.013 (2015).

22 Washburn, E. W. *International Critical Tables of Numerical Data, Physics, Chemistry and Technology*. Vol. 3 54-95 (National Research Council, McGraw-Hill Inc., 1928).

23 Washburn, E. W. *International Critical Tables of Numerical Data, Physics, Chemistry and Technology*. Vol. 2 327-328 (National Research Council, McGraw-Hill Inc., 1928).

24 Ward, W. R. Climatic variations on Mars: 1. Astronomical theory of insolation. *Journal of Geophysical Research* **79**, 3375-3386, doi:10.1029/JC079i024p03375 (1974).

25 Mischna, M. A. On the orbital forcing of Martian water and CO2cycles: A general circulation model study with simplified volatile schemes. *Journal of Geophysical Research* **108**, doi:10.1029/2003je002051 (2003).

26 Forget, F., Haberle, R. M., Montmessin, F., Levrard, B. & Head, J. W. Formation of Glaciers on Mars by Atmospheric Precipitation at High Obliquity. *Science* **311**, 368-371 (2006).

27 Laskar, J. *et al*. Long term evolution and chaotic diffusion of the insolation quantities of Mars. *Icarus* **170**, 343-364, doi:10.1016/j.icarus.2004.04.005 (2004).

28 Nair, H., Allen, M., Anbar, A. D., Yung, Y. L. & Clancy, R. T. A Photochemical Model of the Martian Atmosphere. *Icarus* **111**, 124-150, doi:10.1006/icar.1994.1137 (1994).

29 Gao, P., Hu, R., Robinson, T. D., Li, C. & Yung, Y. L. Stability of CO2 atmospheres on desiccated M dwarf exoplanets *The Astrophysical Journal* **806**, 12, doi:10.1088/0004-637X/806/2/249 (2015).

30 Hanna, J. C. Hydrological modeling of the Martian crust with application to the pressurization of aquifers. *Journal of Geophysical Research* **110**, doi:10.1029/2004je002330 (2005).

31 Fairén, A. G. *et al*. Prime candidate sites for astrobiological exploration through the hydrogeological history of Mars. *Planetary and Space Science* **53**, 1355-1375, doi:10.1016/j.pss.2005.06.007 (2005).

32 Li, D. *et al*. Phase diagrams and thermochemical modeling of salt lake brine systems.II. NaCl+H2O, KCl+H2O, MgCl2+H2O and CaCl2+H2O systems. *CALPHAD: Computer Coupling of Phase Diagrams and Thermochemistry* **53**, 78-89, doi:10.1016/j.calphad.2016.03.007 (2016).

33 Pestova, O. N., Myund, L. A., Khripun, M. K. & Prigaro, A. V. Polythermal Study of the Systems M(ClO4)2-H2O (M2+ = Mg2+, Ca2+, Sr2+, Ba2+). *Russian Journal of Applied Chemistry* **78**, 409-413 (2005).