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)

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