Theory and Evolutionary Evidence of the Autocatalytic Oxygenation of Earth's Surface Environment

by

Haitao Shang

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

Author . . . .
Department of Earth, Atmospheric, and Planetary Sciences
September, 2021

Certified by . . . .
Daniel H. Rothman
Professor of Geophysics
Thesis Supervisor

Accepted by . . . .
Robert van der Hilst
Schlumberger Professor of Earth and Planetary Sciences
Head of Department

Theory and Evolutionary Evidence of the Autocatalytic Oxygenation of Earth's Surface Environment

by

Haitao Shang

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences
on September, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Although molecular oxygen (O_2) is a major component of Earth's atmosphere today and a key signature of life on this planet, we do not understand why and how Earth has evolved from the ancient oxygen-deficient world to the modern oxygen-rich environment, and whether a similar increase can be expected on other planets. This thesis provides a theory to answer this fundamental but unsolved question in Earth science. In the modern environment, atmospheric O_2 is maintained at a stable level due to the existence of negative feedback mechanisms. However, this brings us to a conundrum: under the regulation of negative feedbacks, how could O_2 concentrations have risen? This thesis suggests that the expansion of oxidative metabolisms provided a positive feedback responsible for Earth's oxygenation. This may appear counter-intuitive: oxidative metabolic processes, after all, consume O_2. A potentially important positive feedback nevertheless lies in partially-oxidized organic matter (POOM) produced by oxidative metabolisms in sedimentary environments. This positive feedback derived from oxidative metabolisms is demonstrated via a mathematical model in this thesis. Its relevance to the rise of atmospheric O_2 crucially depends on the existence of POOM-producing oxidative metabolism(s) at the time of Earth's oxygenation(s). One group of enzymes that can catalyze the formation of oxidative metabolic products is the oxygenase family. The methods of molecular phylogenomics are applied to reconstruct the evolutionary history of a representative oxygenase family; the results support such a relevance. Finally, this thesis constructs a mathematical model of Earth's oxygen and carbon cycles and explores the dynamics of these two cycles during oxygenation events. From the perspective of nonlinear dynamics, this mathematical model interprets Earth's oxygenations as dynamical bifurcations of the oxygen cycle and the accompanying excursions in carbon isotope records as the characteristic fluctuations associated with dynamical bifurcations. Collectively, the physical reasoning, phylogenomic analyses, and mathematical modeling in this thesis suggest an unstable evolution of Earth's oxygen and carbon cycles in deep time.

Thesis Supervisor: Daniel H. Rothman
Title: Professor of Geophysics

Acknowledgments

I would like first to thank my advisor, Professor Daniel Rothman. The work presented in this thesis could not have been possible without his endless support. I am especially grateful to the highly free research environment provided by Professor Rothman and also his appropriately nonjudgmental comments on my (sometimes naive) ideas. Professor Rothman's great passion, broad curiosity and unique taste of science, which are what I always wish to emulate, have helped me to develop my own views of scientific research in geoscience; I believe this is the greatest thing that an advisor can do for students.

I thank Professor Tanja Bosak for introducing me to geobiology and teaching me how to think as a scientist. I joined Professor Bosak's lab in 2016 and worked under her guidance for four years. Professor Bosak helped me to become a succinct writer and shrewd reader and gave me lots of support and encouragement during my experimental research.

I also thank Professor Gregory Fournier for introducing me to molecular phylogenetics and supervising my research with great patience. I stepped into this field in Professor Fournier's course in 2017 and had numerous discussion with him in the past three years. I gradually grasped the big picture of this subject under Professor Fournier's “hand-over hand" guidance.

I thank Professor Daniel Repeta and Professor Michael Follows for their genuine interest in my projects, severing in my thesis committee, and the helpful suggestions and comments they gave on my thesis. I also thank Professor Shuhei Ono and Professor Kristin Bergmann for serving as committee members in my qualification exam.

Many thanks to my wonderful friends and colleagues, including Robert Yi, Eric Stansifer, Constantin Arnscheidt, and Yossi Cohen in the Rothman Group, Mirna Daye and Jian Gong in the Bosak Group, Sarah Schwartz and Thiberio Rangel in the Fournier Group, and David Wang and Jeemin Rhim in the Ono Group.

Finally, I am grateful to my family, especially my parents, for their love and support in the past years.

Contents

List of Figures

List of Tables

What do you think is human life like here or there?
It seems like a swan's traces on mud or on snow.
See the claw and nail prints by chance mud and snow bear.
Will the flying swan care what it has left below?
The old monk, dead, has left but a dagoba new;
The verse we wrote was gone with the wall in decay.
What I remember of the journey made with you
Is a weary long way and the lame donkey's bray.

— Su Shi, *Nostalgia – In Response to Ziyou*[*]

---

[*] Yuanchong Xu, *Collected Poems and Lyrics of Classical China.* China Time Books Inc., California, United States (2014).

Chapter 1

Introduction

I am no poet, but if you think for yourselves, as I proceed, the facts will form a poem in your minds.

— Michael Faraday*

Molecular oxygen (O_2) is a remarkable feature of modern Earth and a key signature of life on this planet, but it remains unclear why and how Earth evolved from the ancient O_2-deficient environment to the modern O_2-rich world (Canfield, 2014; Holland, 2006; Lyons et al., 2014; Sessions et al., 2009), and whether a similar evolutionary pattern of oxygen cycle can be expected on other planets such as Mars and Venus. The goal of this thesis is to provide an interpretation for the underlying mechanisms of Earth's oxygenations.

1.1 Essential Concepts

Geochemical studies have suggested that the O_2 concentration in Earth's atmosphere significantly increased during the Great Oxidation Event (GOE), around 2400 to 2300 million years ago (Ma), the Neoproterozoic Oxidation Event (NOE) around 800 to 540 Ma (Canfield, 2014; Holland, 2006; Lyons et al., 2014), and other smaller oxygenation events (Zhang et al., 2018). Figure 1-1 illustrates the evolutionary history of Earth's atmospheric O_2 content.

The escape of hydrogen to outer space has been suggested to be the major source of Earth's atmospheric O_2 during the Archean Eon (Catling et al., 2001a). This source was replaced by oxygenic photosynthesis after the emergence of the great "liberator" – cyanobacteria (Canfield, 2014; Lyons et al., 2014). O_2 and organic matter (CH_2O) are produced in oxygenic photosynthesis:

$$\mathrm{Energy} + H_2O + CO_2 \rightleftharpoons CH_2O + O_2. \tag{1.1}$$

The forward reaction is oxygenic photosynthesis, in which water (H_2O) and carbon dioxide (CO_2) are used to synthesize organic matter, and O_2 is released as a waste product; while the reverse reaction, respiration, burns organic matter with O_2 and provides energy to heterotrophic organisms. Within a short geological time period ($\sim$ 1000 years), almost all of the organic matter produced in the forward reaction is consumed by aerobic respiration or oxidation of reduced compounds and heterotrophic organisms subsequently re-oxidize the vast majority of this organic matter, either directly by O_2 via aerobic respiration or by anaerobic oxidants (such as ferric iron and sulfate) that are

---

*Bence Jones, *The Life and Letters of Faraday.* Longmans, Green and Company, London, United Kingdom (1870).

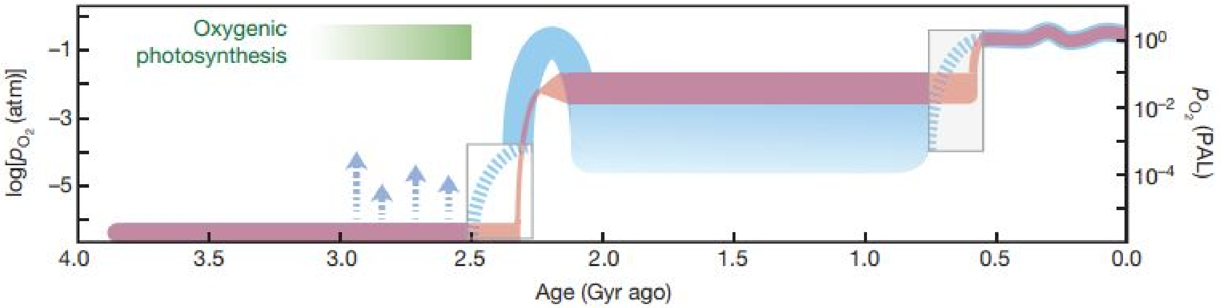


Figure 1-1: Evolutionary history of Earth's atmospheric O_2 content [Lyons et al. (2014); with permission from the Nature Portfolio]. The faded red curve shows a "classical, two-step" view of atmospheric evolution, while the blue curve shows the emerging model (p_{O_2}, atmospheric partial pressure of O_2). Right axis, p_{O_2} relative to the present atmospheric level (PAL); left axis, $\log p_{O_2}$. Arrows denote possible "whiffs" of O_2 late in the Archaean; their duration and magnitude are poorly understood. An additional frontier lies in reconstructing the detailed fabric of "state changes" in atmospheric p_{O_2}, such as occurred at the transitions from the late part of the Archaean to the early Proterozoic and from the late Proterozoic to the early Phanerozoic (blue boxes).

formed through aerobic oxidation by O_2 (Catling and Claire, 2005; Hayes and Waldbauer, 2006; Holland, 1978). A small amount (around 0.1%) of organic matter, however, escapes the respiration and degradation and is eventually buried (i.e., immobilized) into deep sediments, which leads to the accumulation of O_2 in Earth's atmosphere (Catling and Claire, 2005; Hayes and Waldbauer, 2006; Holland, 1978). Most organic matter burial occurs on the continental shelves (Blair and Aller, 2012; Hedges and Keil, 1995; Smith et al., 2015). Although the net O_2 from this "leakage" of carbon cycle seems rather tiny, its effect would be magnified after a long-term ($\sim$ 1,000,000 years) accumulation (Hayes and Waldbauer, 2006; Hedges and Keil, 1995). Atmospheric O_2 is consumed by sink reservoirs: continental oxidative weathering, volcanic reduced gases and metamorphic reduced gases (Catling and Claire, 2005). Oxidative weathering occurs when organic matter is re-exposed to O_2 during mountain uplifts (Canfield, 2014), and volcanic and metamorphic reducing gases (such as H_2, CH_4, H_2S, CO and SO_2), react with O_2 in the atmosphere and seawater (Catling and Claire, 2005; Holland, 1978; Kasting, 2001).

At a basic level, a change in atmospheric O_2 content is expressed as the difference between its source flux F_{source} and its sink flux F_{sink} (Catling and Claire, 2005):

$$\frac{d[\mathrm{O_2}]}{dt} = F_{\text{source}} - F_{\text{sink}}, \tag{1.2}$$

where $[\mathrm{O_2}]$ is the concentration of atmospheric O_2 and t is time. The expression in Equation 1.2 suggests that atmospheric O_2 concentration rises when source exceeds sink (Berner, 2004; Catling and Claire, 2005; Holland, 1978). Under this paradigm, numerous studies have been done to reconstruct the source and sink fluxes (Campbell and Allen, 2008; Duncan and Dasgupta, 2017; Gaillard et al., 2011; Kadoya et al., 2020b; Kump and Barley, 2007; Luo et al., 2016), which envision the rise of atmospheric O_2 as a shift in the equilibrium of the global redox state, following Le Chatelier's principle. In other words, it implicitly assumes that the planetary-scale redox couples on the Earth's surface remain near one steady state; it changes from a low level to a high level during an oxygenation event.

However, an alternative possibility exists: Earth's oxygenation occurs when atmospheric O_2 concentration switches from one equilibrium state to another. This alternative requires the existence of one or more positive feedbacks (Lenton et al., 2008; Prigogine, 1980; Strogatz, 2018). Several studies (Alcott et al., 2019; Goldblatt et al., 2006; Laakso and Schrag, 2017) have suggested some positive feedbacks in Earth's geochemical cycles that may have triggered the transition of atmospheric O_2 steady states. Although purely geochemical feedbacks are possible (Alcott et al., 2019; Goldblatt et al., 2006; Laakso and Schrag, 2017), they may be less responsive to environmental changes compared to the geochemical feedbacks intertwined with biological evolution (Lenton, 1998).

1.2 Core Ideas

The rise of atmospheric O_2 irreversibly changed the chemical environment of Earth's surface and the course of life evolution: anaerobic life was forced to migrate underground (to avoid oxic toxicity), nearly every geochemical cycle was altered by the introduction of oxidative weathering, and microbial ecosystems were fundamentally transformed via the rapid evolution of aerobic metabolisms. The resulting changes in the interaction between life and the environment therefore led to a wholesale re-organization of biogeochemical cycles – that is, a shift from one stable equilibrium to another.

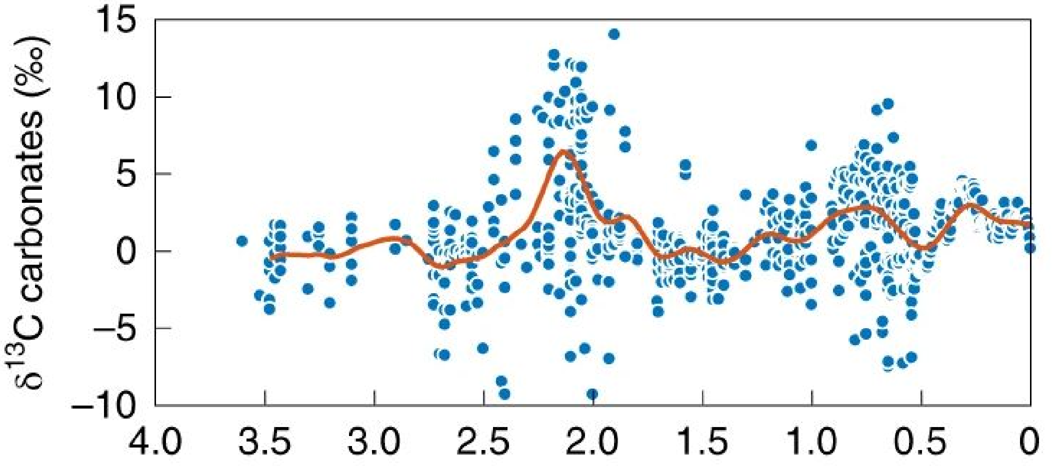


Figure 1-2: Carbonate δ^{13}C through geologic time [Eguchi et al. (2020); with permission from the Nature Portfolio].

In this thesis, I suggest that the requisite positive feedback for Earth's oxygenations derived from the expansion of oxidative metabolisms. This may appear counter-intuitive because oxidative metabolic processes consume O_2. However, the positive feedback lies in the interaction of oxidized metabolic products with minerals in sedimentary environments. In low-O_2 environments, incomplete degradation by oxidative metabolisms results in partially-oxidized organic matter (POOM) that can be persistently protected by minerals. Consequently, POOM may be immobilized in sedimentary rocks, which then enhances the accumulation of O_2 in Earth's atmosphere (Berner, 2004; Catling and Claire, 2005; Holland, 1978).

Based on this positive feedback mechanism, I suggest that the system of Earth's oxygen and carbon cycles undergoes a dynamical bifurcation (Guckenheimer and Holmes, 2013; Prigogine, 1980; Strogatz, 2018) during an oxygenation event. Mathematically, dynamical systems bifurcate when they transition from one equilibrium to another, which involves unstable evolution and is often characterized by significant fluctuations (Guckenheimer and Holmes, 2013; Prigogine, 1980; Strogatz, 2018). The reconfiguration of Earth's biogeochemical cycles suggests that the periods during which O_2 level underwent the greatest changes – the GOE and NOE – were accompanied by the largest known fluctuations in the history of Earth's carbon cycle – the Lomagundi Excursion Event and the Shuram Excursion Event, as shown in Figure 1-2. These observations provide strong support for a shift between alternative equilibrium states during the oxygenation events.

1.3 Content Preview

Chapter 2 presents a hypothesis of partially-oxidized organic matter (POOM) and a positive feedback mechanism it implies. This chapter first constructs a mathematical model of organic degradation according to conventional viewpoints, which suggests a negative feedback that maintains the atmospheric O_2 concentration at a stable level. Next, this chapter introduces the POOM hypothesis and discusses how it is related to the preservation of organic matter. Then, the mathematical model is extended to include the production of POOM; this new model predicts that a positive feedback may exist in low-O_2 environments. How the effect of this positive feedback could be amplified by mineral evolution is also discussed.

Chapter 3 applies the methods of phylogenomics to reconstruct the evolutionary history of one family of oxygenases – Type II Baeyer-Villiger monooxygenases (BVMO) (Landry et al., 2017) – to test the relevance of the POOM hypothesis to Earth's oxygenation(s). The results show that BVMO genes had been acquired very likely at the time of the GOE via a horizontal gene transfer (HGT) event and that the diversification of BVMO genes and the increase in atmospheric O_2 levels are temporally correlated.

Chapter 4 presents a mathematical model of Earth's oxygen and carbon cycles to explore the dynamics of Earth's oxygenation events and carbon isotope disruptions. This model interprets Earth's oxygenation events as dynamical bifurcations and the accompanying fluctuations of carbon isotope as characteristics associated with the bifurcations of Earth's oxygen cycle.

Chapter 5 summarizes the major conclusions of the thesis and provides some closing remarks. Appendix presents the supplementary information.

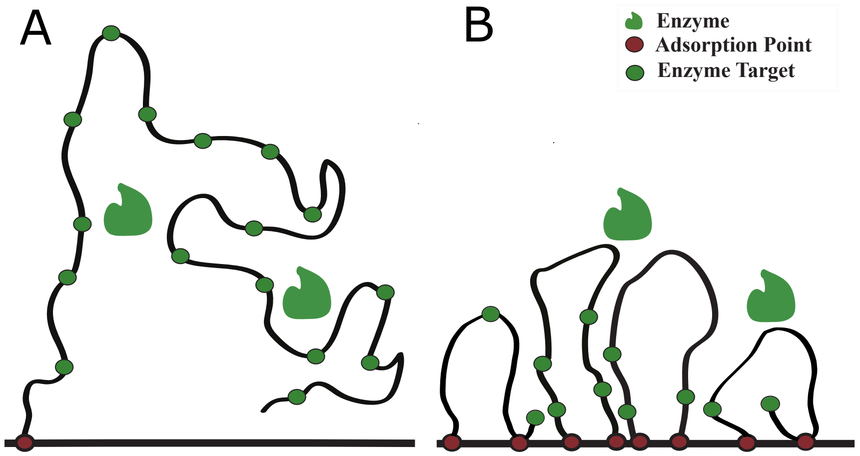
The work in Chapter 2 and Chapter 4 was guided by Professor Daniel Rothman. The work in Chapter 3 was primarily supervised by Professor Gregory Fournier, with some inputs from Professor Daniel Rothman.

Chapter 2

Partially-Oxidized Organic Matter

Every science begins as philosophy and ends as art; it arises in hypothesis and flows into achievement.

— William J. Durant, *The Pleasures of Philosophy*[*]



Abstract

Earth's oxygen and carbon cycles are coupled (Hedges, 2002; Hedges et al., 1999; Hedges and Keil, 1995). Understanding the mechanisms of organic matter degradation and preservation is crucial for deciphering the rise of atmospheric O_2. Instead of presenting a dynamical system of carbon and oxygen cycles to describe the transition between O_2 steady states, this chapter focuses on the conditions under which a positive feedback responsible for Earth oxygenation could have been possible. Field observations have suggested that the modern atmospheric O_2 level is maintained at a steady state due to a negative feedback: higher atmospheric O_2 levels lead to more degradation of organic matter, and consequently O_2 is consumed by the oxidation of organic matter and its level eventually drops back to the initial level (Hartnett et al., 1998; Hedges, 2002). Here I suggest that, in the ancient O_2-deficient environments, the interaction between partially-oxidized organic matter and minerals in sediments can destabilize the steady state(s) of Earth's atmospheric O_2 level, leading to a positive feedback responsible for the accumulation of O_2. This mechanism will be depicted via a mathematical argument in this chapter.

---

[*] William J. Durant, *The Pleasures of Philosophy.* Simon & Schuster, New York, United States (1975).

2.1 Stability of the Modern Atmospheric O_2 Level

The concentration of ambient O_2 is a key factor influencing organic degradation in sediments (Hartnett et al., 1998; Hedges, 2002; Hedges et al., 1999; Hedges and Keil, 1995). (1) O_2 plays the role of electron acceptor in aerobic respiration or the oxidation of reduced compounds (such as sulfide, nitrite, and ferrous iron) (Arndt et al., 2013; Canfield, 1994; Canfield et al., 1993). (2) O_2 can be used as a cofactor by some enzymes (e.g., oxygenases) to degrade organic matter, which is particularly important for nonlabile substances as aromatic compounds (Fuchs et al., 2011). For example, tryptophan 2,3-dioxygenase can insert two O atoms into tryptophan to form N-Formylkynurenine (Nelson et al., 2008), which cleaves an aromatic ring so that further degradation can proceed (Fuchs et al., 2011). If O_2 is absent in the sediments, no equivalent to O_2 can fulfill this function (i.e., being a cofactor of oxygenases) (Emerson and Hedges, 1988; Kristensen et al., 1995). (3) The electron-rich minerals become electron-poor after oxidized by O_2, which can increase the density of positive charges on mineral surfaces and enhance the electrostatic reactions (i.e., Van der Waals force) between positively-charged mineral surfaces and negatively-charged organic molecules (Keil and Mayer, 2014).

Organic matter is highly heterogeneous (Forney and Rothman, 2012; Huang et al., 2003; Keil and Mayer, 2014); so too is the sedimentary environment in which it is deposited (Burdige, 2007; Canfield, 1994; Keil and Mayer, 2014). Compositional and environmental factors both help determine how quickly organic matter decays, and how much is preserved (Arndt et al., 2013; Burdige, 2007; Canfield, 1994; Mayer, 2004). To better understand how mechanisms of preservation influence the stability of O_2 levels, we consider a model of heterogeneity that considers only two types of organic matter: a "labile" component g_1 that ultimately always decays and is never preserved, and a "recalcitrant" component g_2 that decays only in the presence of O_2 (Hedges and Keil, 1995). We assume that, in oxic environments, g_1 degrades with a rate constant k_1 while g_2 degrades with a rate constant $k_2 < k_1$. Organic degradation rate is usually expressed in terms of first-order kinetics (Berner, 1964; Jørgensen, 1978). The aerobic degradation rates at time t are then expressed as

$$\frac{dg_1}{dt} = -k_1 g_1, \tag{2.1}$$

$$\frac{dg_2}{dt} = -k_2 g_2. \tag{2.2}$$

We denote the total amount of organic carbon initially deposited in sediments by $g_0 = g_1(0) + g_2(0)$ and the initial fraction of $g_2(0)$ by a. The initial conditions then read

$$g_1(t = 0) = (1 - a) g_0 \quad \text{and} \quad g_2(0) = a g_0. \tag{2.3}$$

After exposure to O_2 for time t_{ox}, the remaining $g_1(t_{\mathrm{ox}})$ and $g_2(t_{\mathrm{ox}})$ are sequestered in anoxic sediments. According to above assumptions, g_1 will ultimately completely degrade under O_2-free conditions. However, the fate of g_2 is different: its degradation ceases. Burial efficiency – the fraction of organic matter delivered to seafloor that survives subsequent degradation (Hedges and Keil, 1995) – is then

$$\frac{g_2(t_{\mathrm{ox}})}{g_0} = a e^{-k_2 t_{\mathrm{ox}}}, \tag{2.4}$$

Because burial efficiency decreases as oxygen-exposure time (Hartnett et al., 1998; Hedges and Keil, 1995), t_{ox}, increases (Figure 2-1), there is a negative feedback: an increase in O_2 concentration leads to a longer oxygen-exposure time, more O_2 and organic matter are then consumed in degradation, and thus O_2 concentration drops back to its initial level (Hartnett et al., 1998; Rothman, 2015). This negative feedback loop suggests a stability of Earth's oxygen cycle, due to which modern atmospheric O_2 concentration can be maintained at a stable level. However, this brings us to a conundrum: under the regulation of a negative feedback mechanism, how could the O_2 concentration in Earth's atmosphere have risen from the ancient low-level steady state to the modern high-level steady state? Our hypothesis of POOM provides a possible resolution.

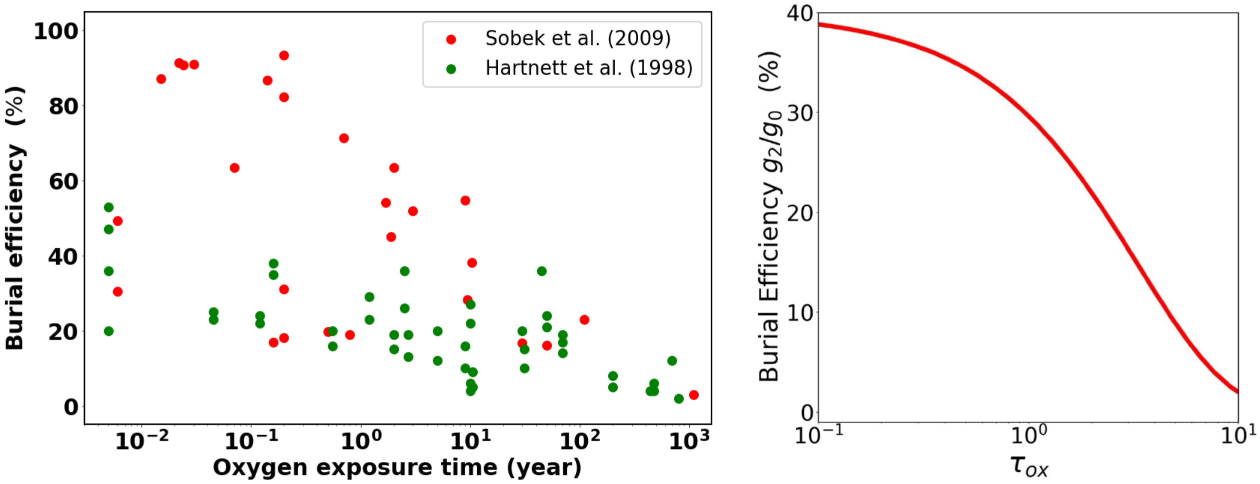


Figure 2-1: Organic carbon burial efficiency as a function of oxygen-exposure time in the modern environment. (a) Field observations; data are from (Sobek et al., 2009) (red dots) and (Hartnett et al., 1998) (green dots). (b) Theoretical predictions. The red curve shows the exponential decay described in (2.4), with $a = 40\%$ and $k_2 = 0.3\,\mathrm{yr}^{-1}$. The horizontal axis represents a dimensionless variable $\tau_{\mathrm{ox}} = k_2 t_{\mathrm{ox}}$.

2.2 A Hypothesis of Partially-Oxidized Organic Matter

Organic degradation is a complex process, which is influenced by the chemical properties of organic matter itself, the physical conditions of surrounding environments, and the biological activities of the organisms in sediments (Arndt et al., 2013; Burdige, 2007; Middelburg and Meysman, 2007; Rothman and Forney, 2007). Microorganisms in sediments secrete enzymes to degrade organic matter, while the association with minerals protects organic matter from enzymatic attack, which is known as "physical protection" (Hemingway et al., 2019; Lalonde et al., 2012; Mayer, 2004). In this section, we suggest that the interaction between intermediate-stage products formed in oxidative metabolisms and minerals in sediments may lead to a positive feedback mechanism.

It has been shown that some oxygen-containing functional groups, such as hydroxyl and carboxyl groups, are abundant in the refractory organic matter in marine (Benner et al., 1992; Cao et al., 2018; Deshmukh et al., 2007; Hay and Myneni, 2007; Hertkorn et al., 2006, 2007; Lechtenfeld et al., 2015; Leenheer, 1994; Nebbioso and Piccolo, 2013; Waggoner et al., 2017), soil (humification) (Baldock and Skjemstad, 2000; DiDonato et al., 2016; Schmidt et al., 2011), and oil reservoirs (Santha et al., 2017; Vandenbroucke and Largeau, 2007; Wu et al., 2012). Here we focus on the organic matter in marine sediments, because most of organic burial occurs on the continental shelves (Blair and Aller, 2012; Hedges and Keil, 1995; Smith et al., 2015).

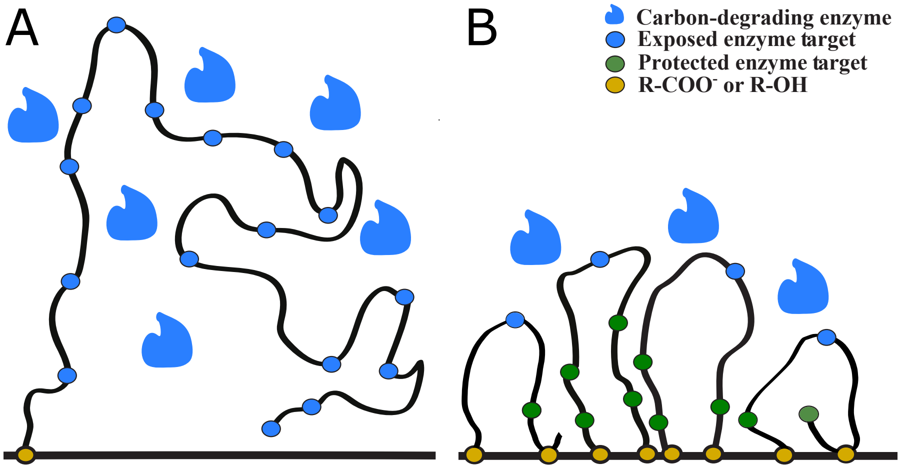


Figure 2-2: Comparison of biopolymers and their interaction with mineral surfaces before (A) and after (B) partial degradation by oxidative metabolisms. In (A), only one site of the biopolymer is sorbed to the mineral surface (horizontal line) and the exposed enzyme targets on the biopolymer are freely accessible to carbon-degrading enzymes secreted by microorganisms. In (B), reactive oxygen-containing functional groups (e.g., R-COO$^-$ or R-OH) are formed by POOM-producing oxidative enzymes such as BVMOs in the presence of O_2. These functional groups create additional sorption sites, enhancing the association of the shorter organic carbon chains with the mineral surface. These partially oxidized, shorter organic carbon chains in (B) are more constrained compared to (A); consequently a large portion of enzyme targets on these shorter organic carbon chains are relatively inaccessible to microbial enzymes. Compared to (A), their degradation requires more investment of free energy to overcome the energy barrier that prevents enzyme access. The juxtaposition of (A) and (B) shows how partial oxidation impedes the biopolymer's accessibility to microbial enzymes and enhances its potential for long-term preservation.

Some oxygen-containing function groups, such as carboxyl groups, have higher adsorption energy – i.e. electrostatic and hydrogen bonding interactions with mineral surfaces – compared to many other common functional groups alkenyl, phenyl, and so on (Ataman et al., 2016; Madsen et al., 1996). An increase in the number of carboxyl groups is able to enhance the adsorption of organic matter on the surface of metal-oxide minerals (Zaouri, 2013). Previous studies have shown that the existence of oxygen-containing functional groups can increase the organic-mineral association in sediments (Schlautman and Morgan, 1994; Tipping, 1981), soil (Kaiser and Guggenberger, 2000; Kile et al., 1999), marine (Drouin et al., 2010) and fresh water (Davis, 1982). Although oxygen-containing functional groups can connect to water molecules through hydrogen bonds, these bonds are weaker than the ionic bonds between the cations on mineral surfaces and the negatively-charged anions on the oxygen-containing functional groups (Keil and Mayer, 2014). Organic matter adsorbed on mineral surfaces can further alter its own three-dimensional orientation via conformational changes and then form more condensed structures to resist microbial enzymes (Baldock and Skjemstad, 2000; DeLapp et al., 2004; Oades, 1988).

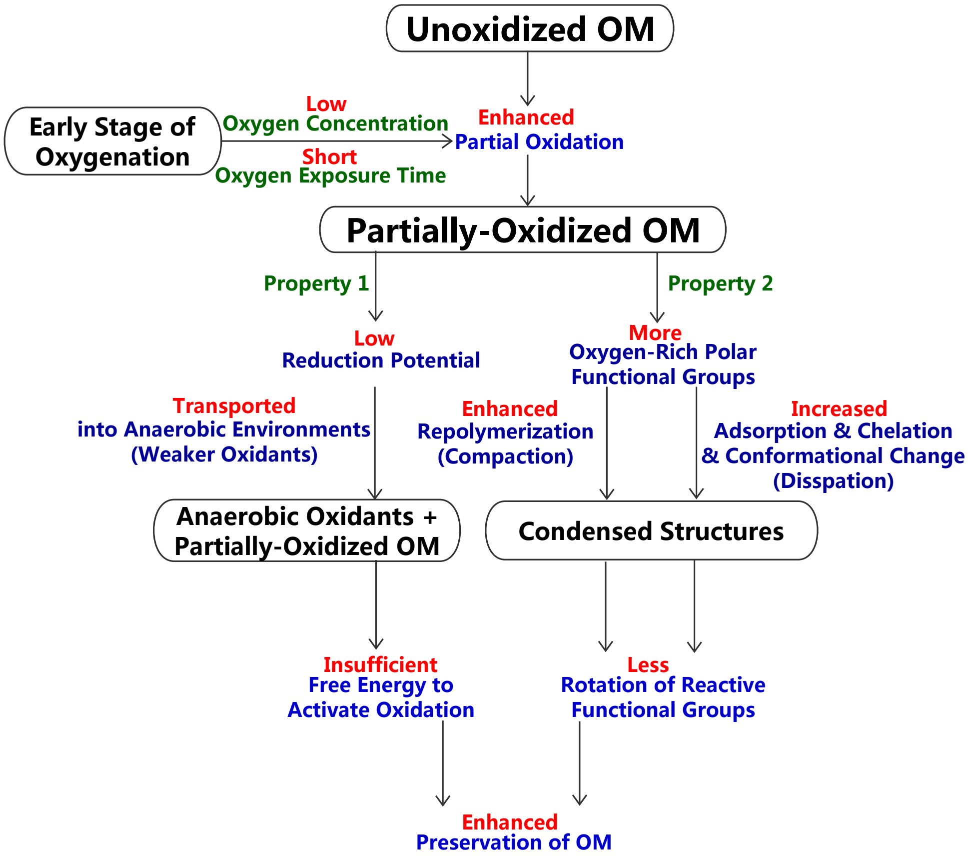


Figure 2-3: A diagram for the mechanisms of POOM hypothesis.

Incomplete degradation of organic matter in the presence of O_2 results in intermediate-stage metabolic products rich in reactive oxygen-containing functional groups such as carboxyls and hydroxyls (Boye et al., 2017; Rashid, 2012). These functional groups have been identified as a characteristic of some recalcitrant organic matter sampled from soil (Wershaw, 1993, 1994) and seawater (Arakawa et al., 2017; Hertkorn et al., 2006, 2007; Repeta, 2015). Laboratory experiments and field studies have demonstrated that reactive oxygen-containing functional groups have high adsorption energy and can be strongly associated with minerals such as iron oxides or clays (Chassé et al., 2015; Frantz et al., 1990; Gu et al., 1994; Rashid et al., 1972; Zhang et al., 2012). According to these observations, we suggest that reactive oxygen-containing functional groups on POOM promotes mineral protection and enhances its recalcitrance. Figure 2-2 provides a conceptual picture. In panel A, enzyme produced by microorganisms can easily access enzyme targets and degrade unoxidized organic matter. In panel B, the clusters of enzyme targets on POOM are beneath a protective shell of less reactive sites and therefore are protected from enzymatic attack, which leads POOM to be persistent in geologic timescale. Based on this POOM model, we hypothesize that, under low-O_2 conditions, the production of POOM rich in reactive oxygen-containing functional groups increases with oxygen-exposure time, which enhances the amount of physically-protected organic matter and promotes organic matter preservation. The mechanisms are summarized in Figure 2-3. However, this positive feedback derived from POOM occurs only when the exposure time to O_2 is brief. Longer exposure to O_2 will act a negative feedback because nearly all organic matter, protected or not, eventually degrades in oxic conditions (Arnarson and Keil, 2007a; Burdige, 2007; Keil and Mayer, 2014; Mayer, 2004).

2.3 A Model of POOM-Involved Organic Degradation

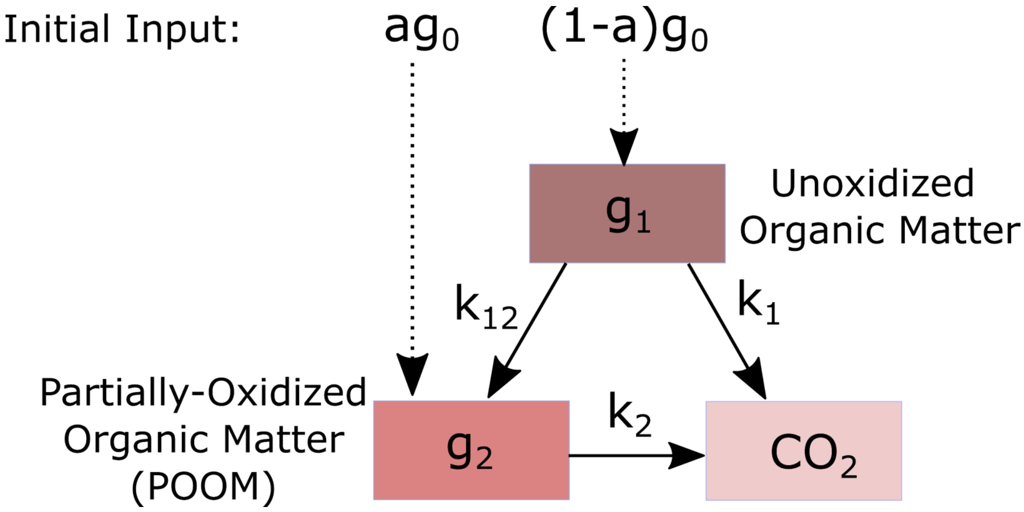


Figure 2-4: Degradation paths of unoxidized organic matter and partially-oxidized organic matter (POOM). g_1 and g_2 represent the amount of unoxidized organic matter and the amount of (physically-protected) POOM deposited in sediments, respectively. Unoxidized organic matter is either directly oxidized to CO_2 with a rate constant k_1 or transformed to POOM with a rate constant k_{12}. POOM is oxidized to CO_2 with a rate constant k_2.

To show how these feedbacks work, we extend (2.1) and (2.2) to include a transformation of unoxidized organic matter (g_1) to POOM (g_2) with a rate constant k_{12}, as shown in Figure 2-5. We assume that (i) organic matter in sediments either is unoxidized (g_1) or has been partially oxidized by O_2 (g_2), (ii) unoxidized organic matter, g_1, is either directly oxidized to CO_2 with a rate constant k_1 or transformed to POOM with a rate constant k_{12}, and (iii) POOM is oxidized to CO_2 with a rate constant k_2 smaller that k_1. The mathematical model reads

$$\frac{dg_1}{dt} = -k_1 g_1 - k_{12} g_1, \tag{2.5}$$

$$\frac{dg_2}{dt} = k_{12} g_1 - k_2 g_2. \tag{2.6}$$

Distinct from the conventional multi-G model (Jørgensen, 1978) and reactive continuum model (Boudreau and Ruddick, 1991) for organic matter degradation, the above mathematical model considers the transformation between organic matter at different degradation stages. According to our assumptions, POOM is either produced in the pre-deposition stage or transformed from unoxidized organic matter in the post-deposition stage. With the reactive oxygen-containing functional groups formed in the presence of O_2, POOM is readily and strongly associated with ambient minerals. After exposed to O_2 for time t_{ox}, remaining unoxidized organic matter $g_1(t_{\mathrm{ox}})$ and POOM $g_2(t_{\mathrm{ox}})$ are sequestered into anoxic sediments, in which biological metabolisms are generally less energetic (Hoehler, 2004; Schrenk et al., 2010). If the free energy yield of the oxidation reaction is smaller than the energy required by microorganisms for ATP production and other biological energy requirements, degradation will cease (Jin and Bethke, 2005; LaRowe and Van Cappellen, 2011). Unoxidized organic matter is generally thermodynamically favorable in the reactions with anaerobic oxidants such as iron(III) and sulfate (LaRowe and Van Cappellen, 2011), thus we assume that eventually $g_1 \to 0$ in anoxic environments. However, POOM has a different fate in anoxic sediments. Intermediate degradation products have been shown to be intrinsically less thermodynamically favorable than unoxidized organic matter (Boye et al., 2017). With the strong protection by minerals, the less energetic POOM, g_2, escapes anoxic degradation and is eventually preserved in deep sediments. Our initial conditions [(2.3)] assume that the total amount of organic matter is unity, therefore $g_2(t_{\mathrm{ox}})$ represents burial efficiency. Together with the initial condition in (2.3), the burial efficiency is then expressed as

$$\frac{g_2(t_{\mathrm{ox}})}{g_0} = \frac{\left[(1-a)e^{(k_2-k_{12}-k_1)t_{\mathrm{ox}}} - 1\right] k_{12} + a(k_2-k_1)}{(k_2-k_1-k_{12})\, e^{k_2 t_{\mathrm{ox}}}}. \tag{2.7}$$

In the above expression of burial efficiency, the degradation rate constants k_1 and k_2 integrate the effect of microbe-mineral competition and other factors influencing organic matter decay. In the degradation continuum (Boye et al., 2017; Jin and Bethke, 2005; LaRowe and Van Cappellen, 2011), POOM produced at different stages may have heterogeneous degradability k_2, which is discussed later in this section. The rate constant k_{12} represents the integrated effect of oxidative metabolism and mineral protection on organic preservation. A positive feedback occurs when $g_2(t_{\mathrm{ox}})/g_0$ increases with t_{ox}; this occurs when

$$k_{12} > \frac{a}{1-a} k_2 \equiv k^{\star}, \tag{2.8}$$

and

$$t_{\mathrm{ox}} < \frac{\log \left[\dfrac{k_2(ak_2 - ak_1 - k_{12})}{(a-1)k_{12}(k_{12}+k_1)} \right]}{k_2 - k_{12} - k_1} \equiv t^{\star}_{\mathrm{ox}}. \tag{2.9}$$

In the degradation continuum (Boye et al., 2017; Jin and Bethke, 2005; LaRowe and Van Cappellen, 2011), POOM produced at different stages may have heterogeneous degradability k_2 (Forney and Rothman, 2012; Rothman and Forney, 2007). Before discussing organic matter degradation and burial efficiency in the degradation continuum, we first introduce $\tau = k_1 t$, $\kappa_{12} = k_{12}/k_1$ and $\kappa_2 = k_2/k_1$ to the equation system of (2.5) and (2.6) and obtain the following dimensionless equation system:

$$\frac{dg_1}{d\tau} = -(1+\kappa_{12})g_1 \tag{2.10}$$

and

$$\frac{dg_2}{d\tau} = -\kappa_{12}g_1 - \kappa_2 g_2. \tag{2.11}$$

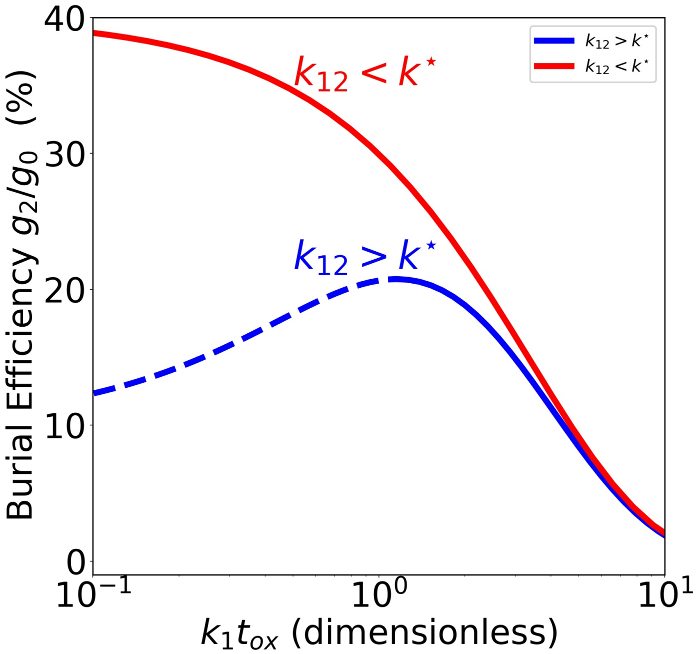


Figure 2-5: Burial efficiency g_2/g_0 as a function of dimensionless oxygen-exposure time $k_1 t_{\text{ox}}$. Positive feedback occurs when burial efficiency increases with oxygen-exposure time (blue dashed line), which appears when $k_{12} > k^\star$ [Equation (2.8)] and $t_{\text{ox}} < t^\star_{\text{ox}}$ [Equation (2.9)].

Substituting the initial conditions in (2.3) into (2.10) and (2.11) gives

$$g_1(\tau) = (1-a)e^{-(1+\kappa_{12})\tau}, \tag{2.12}$$

and

$$g_2(\tau) = \frac{\left[(1-a)e^{(\kappa_2-\kappa_{12}-1)\tau} - 1\right]\kappa_{12} + (\kappa_2 - 1)a}{(\kappa_2 - \kappa_{12} - 1)\, e^{\kappa_2 \tau}}. \tag{2.13}$$

Here, $g_2(\tau_{\text{ox}})$ represents the burial efficiency of total organic matter deposited on seafloor after it has exposed to O_2 for τ_{ox}.

The degradation rate of organic matter depends on both its intrinsic properties and external (that is, physical, chemical and biological) environments (Arndt et al., 2013; Middelburg and Meysman, 2007; Rothman and Forney, 2007). Disordered kinetics (Plonka, 2013; Ross and Vlad, 1999; Vlad et al., 1997) has been suggested as one approach to understand organic matter degradation (Boudreau and Ruddick, 1991). The heterogeneity of degradation rates suggests that there exists a probability density distribution $p(k)$ of rate constant k. It has been proposed that $p(k)$ of mineral-associated organic matter follows: $p(k) \sim k^{-1}$ (Rothman and Forney, 2007). In our model, κ_2 is the degradation rate constant of physically-protected POOM. According to above reasoning, we assume

$$p(\kappa_2) \sim \kappa_2^{-1}, \tag{2.14}$$

where κ_2 is restricted to a finite range $\kappa_{2,\text{min}} \leq \kappa_2 \leq \kappa_{2,\text{max}}$ and $\kappa_{2,\text{min}} \ll \kappa_{2,\text{max}}$. Normalizing $p(\kappa_2)$

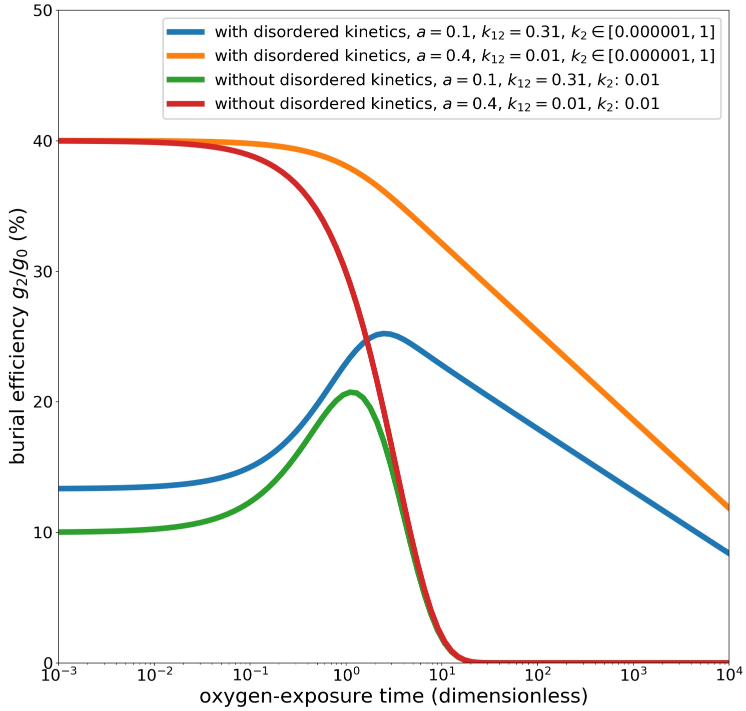


Figure 2-6: Burial efficiency as a function of (dimensionless) oxygen-exposure time with (green and red curves) or without (blue and orange curves) considering the heterogeneity of κ_2 [i.e., Eq. (2.15)]. The hypothesized values for a, κ_{12} and κ_2 are discussed in section 2.4.

over this range gives

$$p(\kappa_2) \sim \begin{cases} \kappa_2^{-1}\left[\log\left(\dfrac{\kappa_{2,\max}}{\kappa_{2,\min}}\right)\right]^{-1}, & \kappa_{2,\min} \le \kappa_2 \le \kappa_{2,\max}, \\ 0, & \text{else.} \end{cases} \tag{2.15}$$

According to our POOM hypothesis and new OET hypothesis, the positive feedback only exists when the oxygen exposure time is short (in a low O_2 environment). In our mathematical model, this means the average burial efficiency $\langle g_2(\tau_{\text{ox}})\rangle$ is an increasing function in the region where $\tau_{\text{ox}} \to 0$. Evaluating the first derivative of $g_2(\tau_{\text{ox}})$ at $\tau_{\text{ox}} = 0$ gives

$$\left.\frac{dg_2}{d\tau_{\text{ox}}}\right|_{\tau_{\text{ox}}=0} = \kappa_{12}(1-a) - \kappa_2 a. \tag{2.16}$$

Thus, we have

$$\left.\frac{d\langle g_2\rangle}{d\tau_{\text{ox}}}\right|_{\tau_{\text{ox}}=0} = \left.\left\langle \frac{dg_2}{d\tau_{\text{ox}}}\right\rangle\right|_{\tau_{\text{ox}}=0} = \kappa_{12}(1-a) - \langle\kappa_2\rangle a, \tag{2.17}$$

where $\langle \kappa_2 \rangle$ is

$$\langle \kappa_2 \rangle = \int_{\kappa_{2,\min}}^{\kappa_{2,\max}} \kappa_2 p(\kappa_2) d\kappa_2 = \frac{\kappa_{2,\max} - \kappa_{2,\min}}{\log(\kappa_{2,\max}/\kappa_{2,\min})}. \tag{2.18}$$

A positive feedback occurs when

$$\kappa_{12} > \frac{a}{1-a} \langle \kappa_2 \rangle \equiv \kappa^\star, \tag{2.19}$$

and

$$\tau_{\text{ox}} < \frac{\log\left[\dfrac{\kappa_2(a\kappa_2 - a - \kappa_{12})}{(a-1)\kappa_{12}(\kappa_{12}+1)}\right]}{\kappa_2 - \kappa_{12} - 1} \equiv \tau_{\text{ox}}^\star. \tag{2.20}$$

The remaining discussion about the parameter space satisfying (2.19) and (2.20) (that is, corresponding to the positive feedback) is the same as the discussion about (2.8) and (2.9) shown in section 2.3. Figure 2-6 shows a comparison of the burial efficiency as a function of oxygen-exposure time when the heterogeneity of κ_2 [that is, disordered kinetics; Eq. (2.15)] is included in the model or not.

The above analyses show that, in the modern O_2-rich environment, burial efficiency monotonically decreases as oxygen-exposure time increases [red curve in Figure 2-5(b)]. When $k_{12} \to 0$, we re-obtain the exponential decay in (2.4). However, when k_{12} is large, a substantial portion of the parameter space [shown by (2.8)] permits a non-monotonic change in burial efficiency when oxygen-exposure time is brief [shown by (2.9)], which is shown by the blue dashed curve in Figure 2-5(b). However, this positive feedback is damped out by a negative feedback after longer exposure to O_2 [blue solid curve in Figure 2-5(b)]. In the ancient low-O_2 environment, the short oxygen-exposure time (due to the shallow penetration depth of O_2) would have led the positive feedback to be ubiquitous.

2.4 Variation of Parameters along the Geologic Time

Previous observations have suggested that the amount of mineral-associated organic matter in marine sediments primarily depends on the sorptive capacity of minerals rather than the supply of organic matter (Keil et al., 1994; Mayer, 1994; Ransom et al., 1998). The sorptive capacity varies with mineral types – that is, mineral physical and chemical properties (Keil and Mayer, 2014; Klaas and Archer, 2002). Since minerals on Earth's surface evolved with geologic time (Hazen and Ferry, 2010; Hazen et al., 2008), we suggest that, a and k_{12}, two parameters in our mathematical model that are related to mineral-protected POOM, changed with mineral evolution and influenced the burial efficiency $g_2(t_{\text{ox}})/g_0$. The mathematical relationship between $g_2(t_{\text{ox}})/g_0$ and these two parameters a and k_{12} is described in (2.7). Based on previous geological and geochemical studies, we assume that the a increases with atmospheric O_2 level while k_{12} decreases with atmospheric O_2 level. The reasoning for this assumption and how the variations of a and k_{12} would have influenced burial efficiency in different geologic times will be discussed in the rest of this section.

Iron(III) oxides (Lalonde et al., 2012; Wagai and Mayer, 2007) is a major mineral type contributing to the long-term preservation of organic matter in modern environments. However, in the pre-GOE world, the majority of iron on Earth's surface existed in a reduced form, either as dissolved Fe(II) or within iron(II)-containing minerals such as pyrite (FeS), and the availability of iron(III) oxides was limited (Anbar, 2008; Holland, 2006). The advent of the GOE dramatically increased the abundance of iron(III) oxides (Hazen et al., 2008; Holland, 2006; Sverjensky and Lee, 2010). The transformation of iron from ferrous to ferric state then could promote the physical protection of organic matter through increasing mineral surface charge density (Baldock and Skjemstad, 2000; Keil and Mayer, 2014) or enhancing the co-precipitation and chelation of organic matter with iron(III) oxides (Barber et al., 2017; Lalonde et al., 2012). As a result, the rise in riverine supply of reactive iron(III) during the GOE increased the amount of pre-protected POOM transported to seafloor – that is, the parameter a. After around 1.5 billion years, the Late Neoproterozoic saw a substantial formation of soils, which hugely promoted the production of clays (Hazen et al., 2008, 2013; Kennedy et al., 2006). Clay minerals have been considered a primary mineral type for the physical protection of organic matter (Blattmann et al., 2019; Kennedy et al., 2002). The rising atmospheric O_2 level in the NOE not only promoted the abundance of iron(III) oxides on Earth's surface, but also facilitated the evolution of multicellular life and therefore enhanced biomineralization (Hazen et al., 2013; Knoll, 2003). Organic matter can be physically protected by biominerals via occlusion into their matrices (Arnarson and Keil, 2007b; Keil and Mayer, 2014). The massive production of clays, iron(III) oxides and biominerals during the Late Neoproterozoic Era again promoted the amount of pre-protected organic matter arriving at the seafloor – that is, the value of parameter a.

The above scenarios of mineral evolution suggest that there should have been two significant increases in the value of a during the GOE and NOE, respectively. How about the other parameter k_{12}? Parameters a and k_{12} are not independent. An increase in the amount of organic matter adsorbed on minerals decreases the density of surface contacts and forms higher repulsive barrier, and therefore lower the adsorption kinetics (Gu et al., 1994; Peterson and Kwei, 1961). As the sorptive load of minerals reaches its capacity, there can no longer be any adsorption. Thus, the production rate constant of physically-protected POOM (that is, k_{12}) should decrease as the amount of pre-formed POOM (that is, a) increases. This implies that the value of k_{12} in our model should have decreased during the transitions of the GOE and NOE.

For convenience, we denote the initial burial efficiency and the transformation rate constant by (i) a_{G} and $k_{12,\mathrm{G}}$ before the GOE, (ii) a_{N} and $k_{12,\mathrm{N}}$ after the GOE but before the NOE, and (iii) a_{M} and $k_{12,\mathrm{M}}$ after the NOE (that is, in modern environment). According to (Fava and Eyring, 1956; Gu et al., 1994), we suggest $k_{12} \sim \exp(-a)$. Thus, for $i, j \in \{\mathrm{G}, \mathrm{N}, \mathrm{M}\}$ and $i \neq j$, we have

$$k_{12,i} \sim (k_{12,j})^{(a_i/a_j)}. \tag{2.21}$$

The results in (Hartnett et al., 1998) suggest that, under modern atmospheric O_2 level, the initial value of average burial efficiency of organic matter is around 0.4 and the range of oxygen-exposure time is from $t_{\mathrm{ox,min}} \sim 10^{-2}$yr to $t_{\mathrm{ox,max}} \sim 10^{4}$yr. Both field observations (Middelburg, 1989; Middelburg and Meysman, 2007) and theoretical analyses (Rothman, 2015; Rothman and Forney, 2007) have suggested that the effective degradation rate constant of organic matter is inversely proportional to its degradation time. Here we propose the following relation between a rate constant k_i and time length under an oxic environment:

$$k_i \sim t_{\mathrm{ox}}^{-1}. \tag{2.22}$$

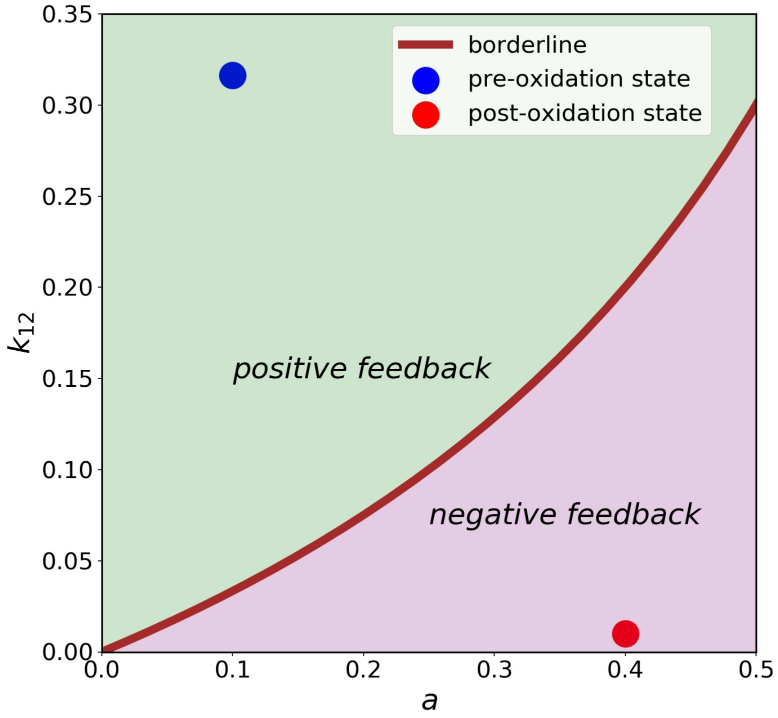


Figure 2-7: The dependence of positive/negative feedback on parameters a and k_{12}. The borderline (brown) for the regions of positive feedback (purple) and negative feedback (pale green) is determined by $k_{12} = k_2 a/(1-a)$ [(2.8)]. The filled blue circle represents a pre-oxidation state; its k_{12} and a values are used for the non-monotonic (blue) curve in Figure 2-5. The filled red circle represents a post-oxidation state; its k_{12} and a values are used for the monotonic (red) curve in Figure 2-5.

Based on the observational data in (Hartnett et al., 1998), we take $g_2|_{t_{\rm ox}=0} = a_{\rm M} \sim 0.4$, $k_{2,\rm min} \sim (t_{\rm ox,max})^{-1} \sim (10^4 \text{yr})^{-1}$ and $k_{2,\rm max} \sim (t_{\rm ox,min})^{-1} \sim (10^{-2}\text{yr})^{-1}$. The fastest degradation rate of POOM should be no faster than the degradation rate of fresh organic matter: $k_{2,\rm max} \leq k_1$. For simplicity, we take $k_1 = k_{2,\rm max} \sim (10^{-2}\text{yr})^{-1}$. Thus, their dimensionless forms are $k_{2,\rm min} = 10^{-6}$ and $k_{2,\rm max} = k_1 = 1$. In the modern environment, the monotonically decreasing g_2 implies $k_{12,\rm M} \lesssim 0.05$. When $k_{12,\rm M} \to 0$, there is effectively no transition from fresh organic matter to POOM, and our model becomes equivalent to the model in (Rothman and Forney, 2007), but with an initial condition a (that is, $g_2|_{t_{\rm ox}=0} = a$) rather than unit. For simplicity, we take $k_{12,\rm M} \sim 0.01$. It has been suggested that the abundance of iron(III) oxides (Holland, 2006) and clay minerals (Kennedy et al., 2002) increased by several times during the transitions of GOE and NOE, respectively. Thus, we assume $a_{\rm N} \sim 0.5 a_{\rm M} \sim 0.2$ and $a_{\rm G} \sim 0.25 a_{\rm M} \sim 0.1$. As discussed above, we have taken $a_{\rm M} \sim 0.4$ and $k_{12,\rm M} \sim 0.01$. Substituting these values into (2.21) gives $k_{12,\rm N} \sim (k_{12,\rm M})^{(a_{\rm N}/a_{\rm M})} \sim 0.1$ and $k_{12,\rm G} \sim (k_{12,\rm M})^{(a_{\rm G}/a_{\rm M})} \sim 0.32$. Since $k_{12,\rm N} > [a_{\rm N}/(1-a_{\rm N})]k_2$ and $k_{12,\rm G} > [a_{\rm G}/(1-a_{\rm G})]k_2$, the positive feedback exits when oxygen-exposure time is short [refer to (2.8)].

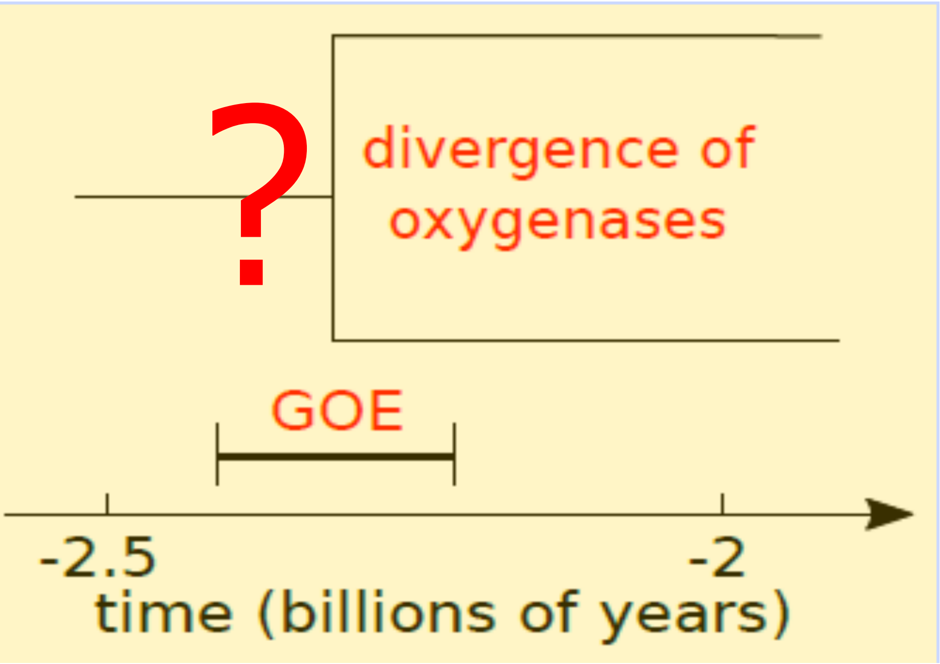
In summary, our POOM hypothesis suggests that the positive feedback responsible for the rise of Earth's atmospheric $\mathrm{O_2}$ levels is a consequence of the combination of biological innovation and geologic changes. In the next chapter, we will apply the methods of molecular phylogenomics to investigate whether POOM really existed at the time of the oxygenation events.

Chapter 3

Molecular Phylogenomics

So we beat on, boats against the current, borne back ceaselessly into the past.

— Scott Fitzgerald, *The Great Gatsby*[*]



Abstract

The POOM hypothesis suggested in Chapter 2 predicts the existence of a positive feedback in the ancient O_2-deficient environment. However, the relevance of the POOM hypothesis to the rise of atmospheric O_2 crucially depends on the existence of POOM-producing oxidative metabolism in the time of Earth's oxygenation. One group of enzymes that can catalyze the formation of intermediate products in oxidative metabolisms is the family of oxygenases. In this chapter, I will apply the methods of phylogenomics to reconstruct the evolutionary history of a relevant oxygenase family and test the relevance of the POOM hypothesis to Earth's oxygenation(s).

---

[*] Scott Fitzgerald, *The Great Gatsby.* Scribner, New York, United States (2004).

3.1 POOM-Producing Genes

The theoretical reasoning and mathematical model in Chapter 2 predict the existence of a positive feedback in low-O_2 environments. However, the relevance of the POOM hypothesis to the rise of atmospheric O_2 crucially depends on the existence of POOM-producing oxidative metabolism(s) in the time of Earth's oxygenation(s). One group of enzymes that can catalyze the formation of intermediate products in oxidative metabolisms is the oxygenase family (Fuchs et al., 2011; Harayama et al., 1992). Both laboratory experiments and field observations have shown that some oxygenases can insert one or two O atoms from O_2 to organic carbon chains to form reactive oxygen-containing functional groups (Fuchs et al., 2011; Harayama et al., 1992; Wershaw, 1993, 1994). Thus, reconstructing the evolutionary history of oxygenases may provide a window through which we can investigate whether and how oxygenases are responsible for Earth's oxygenation. The information from geologic records, however, cannot directly reveal the evolution of enzymes, but the methods of phylogenomics can be used to reconstruct and analyze the genomic information in extinct ancient species (Kacar et al., 2017; Yang, 2014). We will apply these methods to test the relevance of POOM to Earth's oxygenation in this chapter.

The relevance of the POOM hypothesis to the rise of atmospheric O_2 crucially depends on the existence of POOM-producing oxidative metabolisms during Earth's oxygenation. One group of enzymes that can catalyze the formation of intermediate products in oxidative metabolisms are oxygenases (Harayama et al., 1992). Both laboratory experiments and field observations have shown that some oxygenases can insert one or two O atoms from O_2 into organic carbon chains to form reactive oxygen-containing functional groups (Harayama et al., 1992; Wershaw, 1993). Thus, reconstructing the evolutionary history of relevant oxygenases may provide a window through which we can investigate the possible role of POOM-producing oxidative metabolisms in Earth's oxygenation. The information from geologic records cannot directly reveal the evolution of enzymes, but phylogenetic methods can be used to reconstruct and analyze the history of enzyme families across geological timescales (Yang, 2014).

SAR202 bacteria (members of the phylum Chloroflexi) are ubiquitous in various modern environments and account for a significant fraction of the microbial population in the deep ocean (Mehrshad et al., 2018; Morris et al., 2004). They are especially prevalent in marine environments where recalcitrant organic matter is abundant (Landry et al., 2017; Liu et al., 2020). Landry et al. (2017) have demonstrated that SAR202 bacteria produce recalcitrant deep-ocean organic matter containing reactive oxygen-rich functional groups such as hydroxyls and carboxyls; moreover, they show that the Baeyer-Villiger monooxygenase (BVMO) enzymes play a predominant role in these oxidative metabolisms. BVMOs are a family of oxygenases that can catalyze a wide variety of oxidation reactions of a large range of substrates (Tolmie et al., 2019). Here, we reconstruct the evolutionary history of BVMOs in SAR202 to test the hypothesis that POOM-producing oxidative metabolisms and Earth's oxygenation are temporally correlated.

3.2 Phylogenies Reconstruction

Protein Basic Local Alignment Search Tool (BLASTp) on the National Center for Biotechnology Information (NCBI) database was used to search the genes of interest. For the species tree reconstruction, we used the sequences that are homologous to 30 ribosomal proteins (Table A.1) of

SAR202 cluster bacterium Io17-Chloro-G4. Ribosomal proteins are from 298 taxa Table A.2, including 203 taxa directly collected from the NCBI database and 95 taxa provided in (Magnabosco et al., 2018). To reconstruct the gene tree of BVMOs, we used 330 sequences homologous to the BVMO of *SAR202 cluster bacterium* Io17-Chloro-G4 (NCBI Query ID: PKB68843.1). To check the homology of those oxygenase genes that were not well annotated in the original database, we employed the method used in (Landry et al., 2017), manually ran these sequences on Phyre2 (Kelley et al., 2015) and confirmed that they have high predicted structural similarity to BVMOs. The program MAFFT (Katoh and Standley, 2013) was used to align the sequences. Alignments were visualized on Clustal X (Larkin et al., 2007), and poorly aligned regions were manually deleted. The alignments were concatnated using the program SequenceMatrix (Vaidya et al., 2011).

Both the maximum-likelihood species tree and gene tree were inferred using IQ-TREE (version 1.6.3) (Nguyen et al., 2015) with the ultrafast bootstrap algorithm (Hoang et al., 2018) and 100 replicates. The ModelFinder Plus option was used in IQ-TREE to identify the best-fit substitution model for each sequence dataset. The species tree alignment best fit the “LG+I+R9” model, using an LG model for amino-acid exchange rate matrices (Le and Gascuel, 2008), invariable sites, and a Free Site Rate model with 9 categories of Gamma-distributed rates (Soubrier et al., 2012). The “LG+R6” model best fit the BMVO gene tree, using an LG model with 6 Free Site Rate categories.

Previous studies and outgroup rootings have consistently supported a species tree rooting of Bacteria where Proteobacteria, Bacteroidetes, Ignavibacteria and Chlorobi are grouped in one clade, while Actinobacteria, Chloroflexi and Cyanobacteria are grouped in the other clade (Battistuzzi and Hedges, 2009). Therefore, we manually rooted the species tree based on above information. Since we have no clear knowledge about the root for the gene tree of BVMOs, the minimal ancestor deviation (MAD) method (Tria et al., 2017) was used to infer its root position. The reconstructed rooted maximum likelihood species tree of BVMO-containing taxa is presented in Figure 3-1, and the reconstructed rooted maximum likelihood tree of BVMO homologs is presented in Figure 3-2.

3.3 Divergence Time Estimation

The divergence time was inferred using the program PhyloBayes (Version 4.1c) (Lartillot et al., 2009) with "-catfix C60 -ugam -bd -sb -nchain 2 100 0.3 50" and all other parameters default. The timing of the earliest evidence for Earth’s habitability – 4.4 Ga zircon (Valley et al., 2014) – was used as the older (i.e., lower) bound for the root prior, while the timing of the earliest known bacterial microfossils – 3.4 Ga stromatolites in Warrawoona Group, Australia (Schopf, 1993) was used as the younger (i.e., upper) bound for the root prior. Since our belief in the true root age becomes weaker as the date approaches the two bounds (4.4 Ga and 3.4 Ga), we approximated this by using a normally distributed root prior with a 95% confidence interval across the range (that is, a mean of 3.9 Ga and a standard deviation of 0.25 Ga). The uncorrelated gamma model (i.e., -ugam) (Lemey et al., 2009) was used for the relaxed molecular clock. Since the estimated oldest habitability was 4.4 Ga (Valley et al., 2014) and we have a strong belief that the node of last universal common ancestor (LUCA) should be deeper than the root node on the species tree (Weiss et al., 2016), a hard bound was imposed on the older age of the root prior in the calibration file. However, we set a soft bound (Yang, 2014) on the younger age of the root prior. While it is unlikely that the 3.4 Ga stromatolites represent stem bacterial diversity (Magnabosco et al., 2018), it remains a possibility, and therefore the model should permit the root age of the species tree to be younger than 3.4 Ga.

The 95% confidence intervals of posterior ages on following crown and stem nodes obtained in previous study (Magnabosco et al., 2018) were used as the calibrations for the species tree: (1) Crown Chloroflexi: 3.374 Ga – 2.674 Ga; (2) Stem Green Non-sulfur Bacteria: 2.782 Ga – 2.053 Ga; (3) Crown Green Non-sulfur Bacteria: 2.041 Ga – 1.645 Ga; (4) the common ancestor of Ktedonobacteria and Dehalococcoidia: 2.995 Ga – 1.793 Ga; (5) Crown Dehalococcoidia: 2.013 Ga – 0.789 Ga; (6) the common ancestor of Anaerolineae and Green Non-sulfur Bacteria: 3.124 Ga - 2.384 Ga. These uniform age priors have broad ranges and therefore are unlikely over-specifying due to some bias from previous analyses (Magnabosco et al., 2018), which is appropriate because our knowledge of the divergence time on these nodes is limited. The prior ages for the species tree are listed in Table 3.1.

The chronogram of species tree was generated using the "-readdiv" command in PhyloBayes (Lartillot et al., 2009). Since PhyloBayes is a phylogenetic Monte Carlo Markov Chain (MCMC) sampler (Lartillot et al., 2009), the sampled values in the initial phase are likely to be outside a high probability region (Lemey et al., 2009; Yang, 2014). To exclude the influence of the starting point on chronogram computation, the first 20% of sampled node ages were discarded (i.e., "burned in"). The command "-v" in PhyloBayes was used to output all dated trees of the samples in the “datedist" file, which were used to calculate the node age distributions and the divergence rates of genes. To compare posterior ages to estimations obtained without the information from gene sequences, we generated the age estimates using the "-prior" command in PhyloBayes.

The 95% confidence intervals for five nodes in the SAR202 clade on the species tree were used as the secondary calibrations to estimate the divergence time on the gene tree. These SAR202 nodes in the species tree all have high bootstrap supports ($> 90\%$) and are also represented in the gene tree. It has been suggested that the diversification time of oxygen-consuming metabolisms should be roughly consistent with the timing of the GOE (Kim et al., 2012). However, due to the lack of prior knowledge about the root of the gene tree, we imposed a generous flat root prior of 3.2 Ga - 1.2 Ga to avoid false precision. The older bound of this range (i.e., 3.2 Ga) not only corresponds to the published age estimates (Magnabosco et al., 2018) for stem Cyanobacteria (i.e., the split between Cyanobacteria and non-photosynthetic lineages), but also is older than nearly all proposed geochemical evidence for oxygen (Anbar et al., 2007; Lyons et al., 2014). The younger bound was set as 1.2 Ga, which is sufficiently young to include the 95% confidence interval for the age of the crown node of the SAR202 clade on the species tree, from which we expected to use secondary calibrations for the gene tree. The prior ages for the gene tree are listed in Table 3.1.

The chronogram for the species tree of BVMO-containing taxa is presented in Figure 3-3, and the chronogram for the gene tree of BVMO homologs is presented in Figure 3-4. Figure 3-7-A shows a subtree of the chronogram for SAR202 clade and related Dehalococcoidia group. The posterior date distribution of the stem SAR202 node (red) in Figure 3-7-A suggests that an ancestral SAR202 species likely had existed at the time of the GOE.

Figure 3-1: Rooted maximum likelihood species tree of BVMO-containing taxa, including the SAR202 cluster within Chloroflexi. Bipartitions are labeled with percentage support values (approximate likelihood ratio/bootstrap). This tree is manually rooted so that Proteobacteria, Bacteroidetes, Ignavibacteria and Chlorobi are grouped in one clade while Actinobacteria, Chloroflexi and Cyanobacteria are grouped in the other clade.

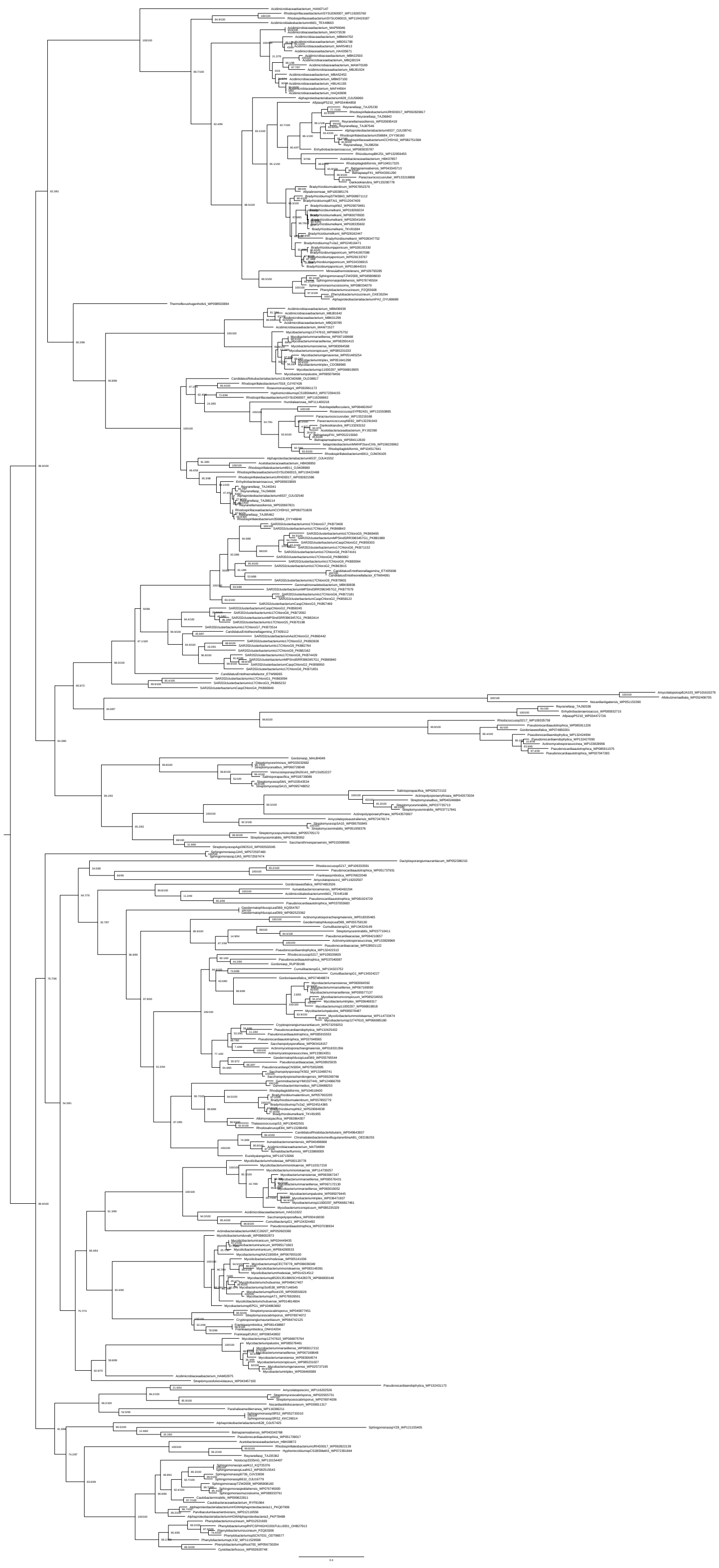


Figure 3-2: Rooted maximum likelihood tree of BVMO homologs. Bipartitions are labeled with percentage support values (approximate likelihood ratio/bootstrap). This tree is rooted using the minimal ancestor deviation (MAD) method (Tria et al., 2017).

Table 3.1: Secondary Age Calibrations for Species and Gene Trees

Phylogeny	Node	Upper limit (Ga)	Lower limit (Ga)
Species Tree	Root	4.400	-1
	Crown Chloroflexi	3.374	2.674
	Stem Chloroflexia	2.782	2.053
	Crown Chloroflexia	2.041	1.645
	Common Ancestor of Classes Ktedonobacteria and Dehalococcoidia	2.995	1.793
	Crown Dehalococcoidia	2.013	0.789
	Common Ancestor of Classes Anaerolineae and Chloroflexia	3.124	2.384
Gene Tree	Root	3.200	1.200
	Common Ancestor of *SAR202 Cluster Bacterium* Io17-Chloro-G1 and Casp-Chloro-G4	1.623	0.988
	Common Ancestor of *SAR202 Cluster Bacterium* Io17-Chloro-G6 and Io17-Chloro-G8	0.620	0.358
	Common Ancestor of *SAR202 Cluster Bacterium* Io17-Chloro-G2 and Io17-Chloro-G9	0.587	0.265
	Common Ancestor of *SAR202 Cluster Bacterium* Io17-Chloro-G7 and Casp-Chloro-G2	1.083	0.721
	Common Ancestor of *SAR202 Cluster Bacterium* Io17-Chloro-G6 and Casp-Chloro-G3	1.343	0.891

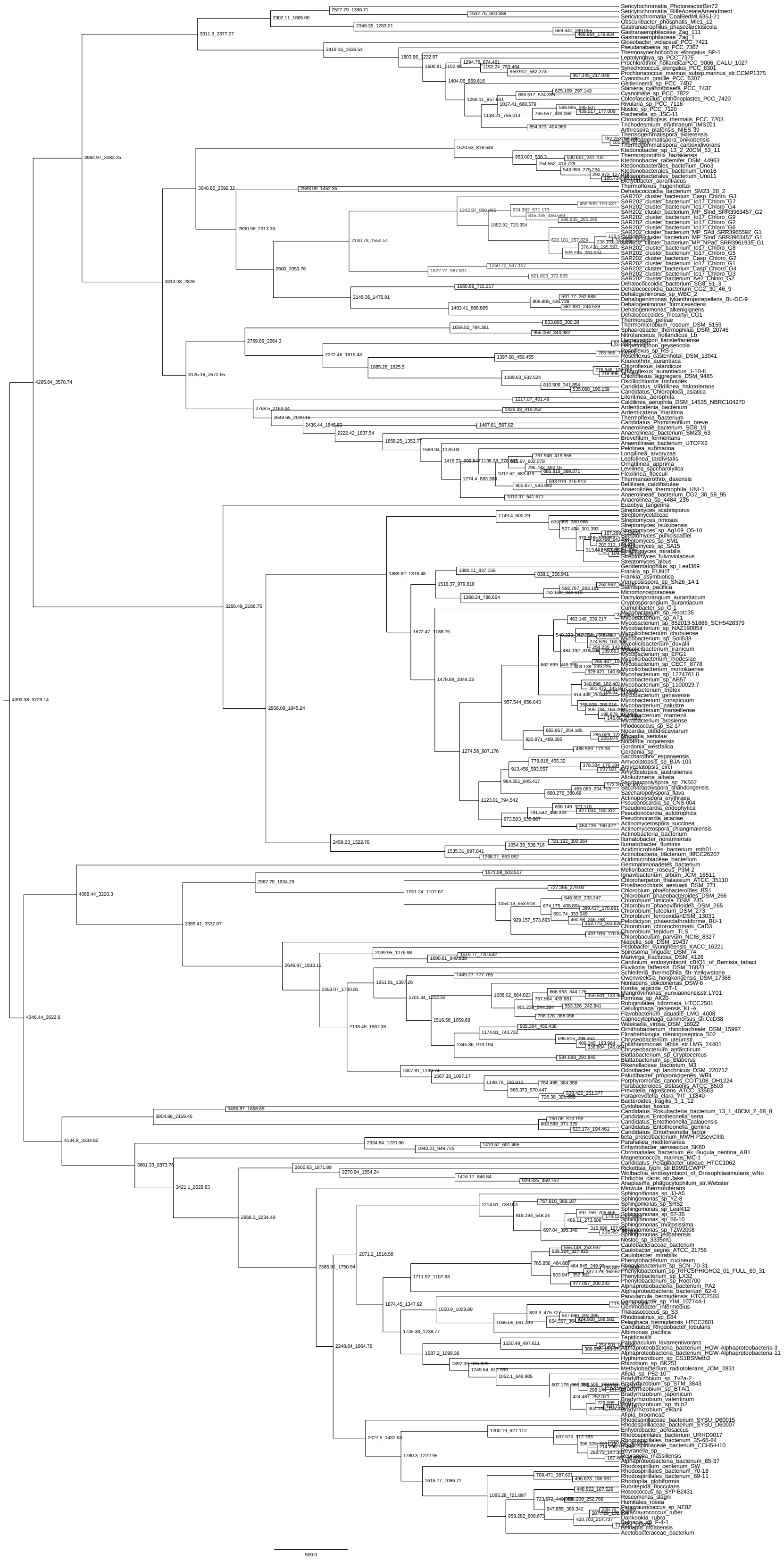


Figure 3-3: Species tree chronogram of BVMO-containing taxa, including the SAR202 cluster within Chloroflexi. Values on nodes represent 95% confidence intervals of inferred posterior age distributions.

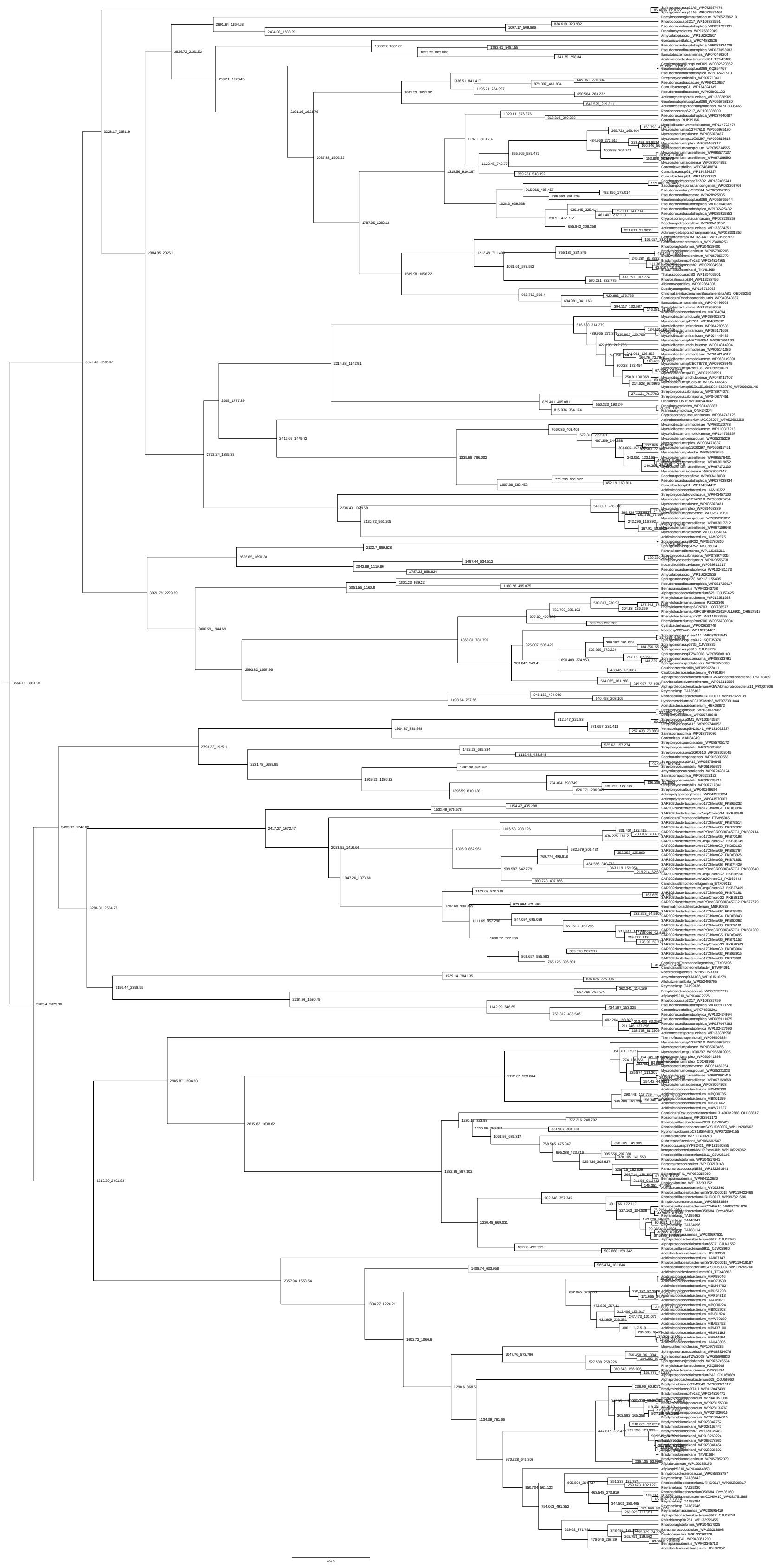


Figure 3-4: Gene tree chronogram of BVMO homologs. Values on nodes represent 95% confidence intervals of inferred posterior age distributions.

Table 3.2: The 95% Confidence Intervals for Prior and Posterior Age Values of Representative Nodes on Species Tree

Node	Prior	Posterior
Root of Species Tree	4.382 Ga – 3.701 Ga	4.393 Ga – 3.729 Ga
Crown *Chloroflexi*	3.364 Ga – 2.848 Ga	3.313 Ga – 2.828 Ga
Stem *Green Non-Sulfur Bacteria*	2.751 Ga – 2.234 Ga	2.770 Ga – 2.264 Ga
Crown *Green Non-Sulfur Bacteria*	1.888 Ga – 1.629 Ga	1.885 Ga – 1.626 Ga
Common Ancestor of Classes *Ktedonobacteria* and *Dehalococcoidia*	3.052 Ga – 2.599 Ga	3.041 Ga – 2.592 Ga
Crown *Dehalococcoidia*	1.559 Ga – 0.949 Ga	1.483 Ga – 0.867 Ga
Common Ancestor of Classes *Anaerolineae* and *Green Non-Sulfur Bacteria*	3.129 Ga – 2.671 Ga	3.125 Ga – 2.672 Ga
Stem *Cyanobacteria*	3.335 Ga – 2.461 Ga	3.311 Ga – 2.377 Ga
Crown *Cyanobacteria*	2.618 Ga – 1.792 Ga	2.419 Ga – 1.637 Ga
Crown *Cyanobacteria* (excluding *Geobacter*)	2.076 Ga – 1.449 Ga	1.804 Ga – 1.233 Ga
Stem *Green Sulfur Bacteria*	3.027 Ga – 1.904 Ga	2.983 Ga – 1.934 Ga
Crown *Green Sulfur Bacteria*	2.063 Ga – 1.229 Ga	1.901 Ga – 1.108 Ga

3.4 Horizontal Gene Transfer

The synergetic interactions between environmental changes and biological innovations probably result in horizontal gene transfer (HGT) events (Fournier et al., 2009; Soucy et al., 2015) of oxygenases at the time of Earth oxidations. HGT events of BOVM genes were inferred by reconciling the gene tree with the species tree using RANGER-DTL (Bansal et al., 2018). Phylogenetic reconciliation was performed using the program RANGER-DTL (Version 2.0) (Bansal et al., 2018). The cost scores of duplication, transfer, and loss (DTL) were set as default values (i.e., 2, 3 and 1). The rooted species (Figure 3-5) and the rooted gene tree (Figure 3-6) are reconciled to infer the horizontal gene transfer (HGT) events occurred in the evolutionary history of BVMOs. The reconciliation was repeated 100 times, and the estimated HGT events that have high bootstrap supports ($\geq 80\%$) were selected for further analyses.

To investigate the relevance of these HGT events to evolution of oxygen and carbon cycles, we constructed the weighted distributions of the older and younger bounds for the timing of 68 HGT acquisitions that have "$\geq 80\%$" bootstrap values. The age distributions of the recipients for these HGT events were computed using the "datedist" file of species tree and the following rules: (i) for transfers that have multiple (more than one) non-leaf recipients, we computed the fraction of each recipient as its weight and used these weights to calculate the weighted age distribution; (ii) for HGTs which have one or more than one leaf recipients, we deleted the leaf recipients, only kept the "internal node" recipients, computed the fraction of each recipient as its weight and used these weights to calculate the weighted age distribution; (iii) for HGTs that have one or more than one leaf donors, we included the ages of the recipients for these transfers when constructing the age distributions. These weighted age distributions of recipients show the ranges of the older and younger bounds for the inferred HGT events, which is graphically summarized in Figure 3-7.

Among the inferred HGT events (Figure 3-7), the earliest acquisition of BVMO gene (i.e., HGT event $\#1$) occurred on the branch between the stem and crown SAR202 nodes [Figure 3-7, Inset A]. The older and younger bounds for the timing of this HGT event therefore correspond to the ages of the stem and crown SAR202 nodes, respectively. There is a 95% probability of this initial HGT into SAR202 occurring between 2600 Ma and 1540 Ma, with 85% and 97% of the probability density of the donor-recipient branch interval intersecting the GOE and the Lomagundi Excursion Event, respectively [Figure 3-7, Inset B]. The Lomagundi Excursion Event, the largest known positive carbon isotope excursion event in Earth's history, has been interpreted as a consequence of a dramatic increase in the burial of organic carbon (Martin et al., 2013). The temporal overlap of the initial HGT acquisition of BVMO gene with the GOE and the Lomagundi Excursion Event suggests that there may exist some linkages between the diversification and ecological dispersal of POOM-producing genes and these significant geological events. Figure 3-7 also shows that the HGT events of BVMO genes span the Proterozoic and Phanerozoic, apparently increasing in frequency starting in the late Neoproterozoic. However, the relatively large uncertainties in the dating of these HGT events prevent a more specific interpretation of whether and how this increasing trend of HGT frequency is related to the higher-resolution histories of oxygen and carbon cycles and other geological events, especially in the Phanerozoic, although such relations may exist. Future improvements in molecular clock methodologies may increase the precision of these estimates, enabling more detailed hypothesis testing.

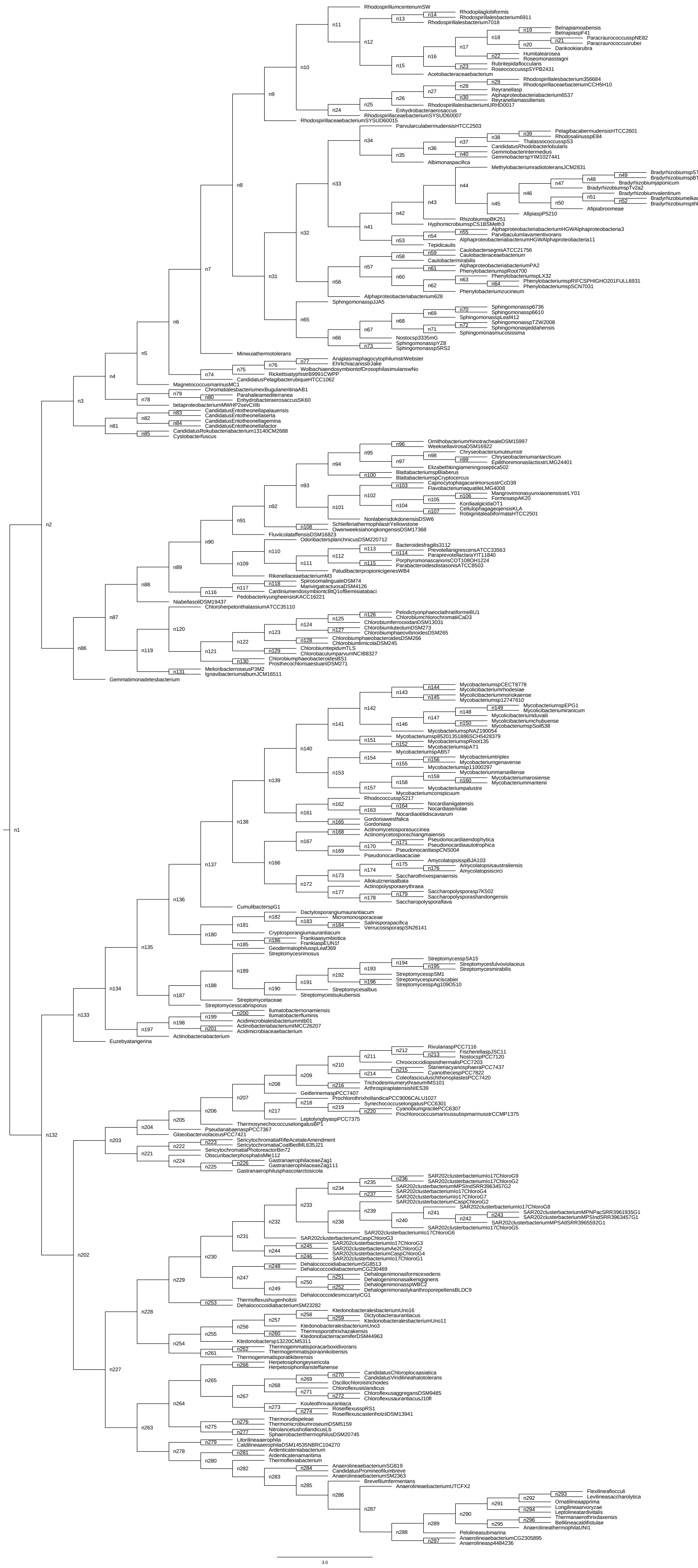


Figure 3-5: Rooted species tree cladogram with RANGER-DTL nodes mapped (n#).

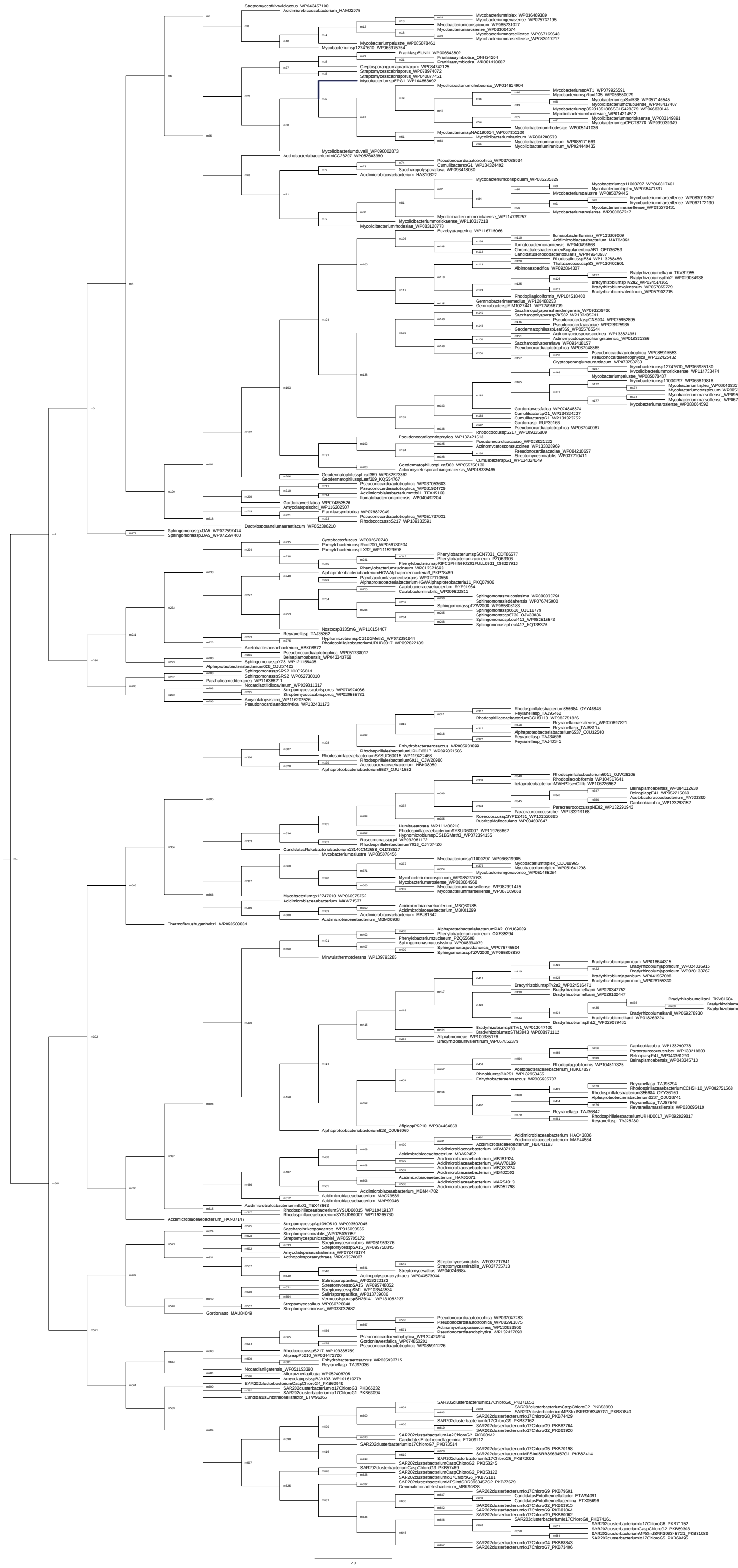


Figure 3-6: Rooted BVMO homolog tree cladogram with RANGER-DTL nodes mapped (m#).

Table 3.3: Weighted Means and 95% CIs of the Older and Younger Bounds for HGT Events

(Refer to Figure 3-5 and Figure 3-6 for the information of node labels "n#" and "m#".)

HGT ID	Node in Gene Tree	Recipient Branch in Species Tree (older bound, younger bound)	Recipient Weights	Weighted Mean Ages (Ma) (older bound, younger bound)	95% CI (Ma) (older bound, younger bound)
1	m561	(n230, n231)	100 %	(2350.04, 1830.44)	(2598.23 - 2056.45, 2110.15 - 1535.13)
2	m3	(n134, n135)	23 %	(2002.70, 1286.83)	(2407.04 - 1616.46, 1552.85 - 1054.73)
		(n136, n180)	11 %		
		(n197, n198)	66 %		
3	m398	(n32, n56)	100 %	(1595.38, 1421.68)	(1847.61 - 1341.72, 1703.12 - 1130.51)
4	m397	(n32, n56)	100 %	(1595.38, 1421.68)	(1847.61 - 1341.72, 1703.12 - 1130.51)
5	m101	(n135, n136)	73 %	(1465.22, 1268.09)	(1722.05 - 1252.46, 1471.50 - 1078.70)
		(n138, n166)	13 %		
		(n198, n199)	14 %		
6	m450	(n9, n10)	65 %	(1426.47, 1148.75)	(1659.50 - 1203.30, 1363.03 - 955.47)
		(n12, n15)	10 %		
		(n24, n25)	25 %		
7	m27	(n136, n180)	100 %	(1420.51, 1252.02)	(1667.97 - 1201.51, 1506.32 - 984.37)
8	m209	(n198, n199)	100 %	(1188.4, 770.51)	(1570.97 - 892.80, 1068.74 - 583.46)
9	m210	(n198, n199)	100 %	(1188.4, 770.51)	(1570.97 - 892.80, 1068.74 - 583.46)
10	m118	(n34, n35)	87 %	(1106.39, 761.17)	(1320.90 - 907.90, 942.82 - 603.34)
		(n50, n51)	13 %		
11	m452	(n12, n15)	100 %	(892.13, 764.41)	(1086.88 - 723.66, 957.47 - 608.02)
12	m563	(n24, n25)	100 %	(889.7, 440.02)	(1199.50 - 645.63, 645.03 - 317.33)
13	m579	(n24, n25)	100 %	(889.7, 440.02)	(1199.50 - 645.63, 645.03 - 317.33)
14	m626	(n233, n238)	100 %	(883.07, 473.44)	(1072.91 - 722.48, 617.93 - 358.03)
15	m25	(n147, n148)	38 %	(861.00, 719.55)	(1046.02 - 708.32, 898.87 - 566.69)
		(n136, n180)	19 %		
		(n198, n201)	43 %		
16	m413	(n44, n45)	100 %	(838.21, 472.71)	(1050.19 - 662.94, 600.62 - 355.46)
17	m584	(n172, n173)	100 %	(800.92, 748.67)	(967.57 - 648.89, 920.87 - 593.96)
18	m150	(n167, n168)	100 %	(797.08, 281.58)	(975.44 - 640.32, 448.32 - 161.99)
19	m144	(n167, n169)	100 %	(797.08, 619.41)	(975.44 - 640.32, 796.96 - 467.36)
20	m301	(n139, n161)	100 %	(794.51, 649.30)	(961.42 - 652.49, 818.11 - 485.25)
21	m104	(n35, n36)	76 %	(793.82, 588.95)	(985.13 - 629.09, 757.10 -452.10)
		(n36, n37)	24 %		
22	m108	(n199, n200)	100 %	(770.51, 418.98)	(1068.74 - 538.46, 728.29)
23	m584	(n252, n253)	100 %	(735.50, 624.62)	(924.00 - 568.82, 807.58 - 466.02)
24	m339	(n13, n14)	100 %	(574.24, 310.48)	(789.73 - 388.53, 493.71 - 185.01)
25	m254	(n57, n58)	50 %	(562.78, 419.13)	(687.26 - 453.97, 532.49 - 329.50)
		(n67, n68)	50 %		

Continued

HGT ID	Node in Gene Tree	Recipient Branch in Species Tree (older bound, younger bound)	Recipient Weights	Weighted Mean Ages (Ma) (older bound, younger bound)	95% CI (Ma) (older bound, younger bound)
26	m637	(n82, n84)	100 %	(562.30, 336.65)	(845.71 - 369.75, 561.40 - 197.73)
27	m645	(n234, n237)	90 %	(558.50, 279.98)	(688.41 - 446.07, 445.74 - 167.27)
		(n239, n240)	10 %		
28	m333	(n12, n13)	5 %	(552.2, 377.67)	(686.45 - 442.86, 501.16 - 274.61)
		(n17, n22)	95 %		
29	m232	(n25, n26)	84 %	(496.72, 317.45)	(672.75 - 385.20, 418.62 - 248.31)
		(n56, n57)	4 %		
		(n57, n60)	2 %		
		(n66, n67)	10 %		
30	m336	(n13, n14)	7 %	(470.75, 255.35)	(581.22 - 377.09, 357.08 - 176.56)
		(n20, n21)	28 %		
		(n16, n23)	65 %		
31	m140	(n167, n169)	31 %	(466.12, 256.36)	(596.10 - 364.73, 329.92 - 197.61)
		(n178, n179)	69 %		
32	m337	(n13, n14)	17 %	(460.79, 251.41)	(566.36 - 368.23, 343.86 - 179.92)
		(n13, n14)	30 %		
		(n13, n14)	53 %		
33	m564	(n170, n171)	100 %	(445.86, 286.04)	(601.87 - 323.96, 419.72 - 178.85)
34	m157	(n170, n171)	100 %	(445.86, 286.04)	(601.87 - 323.96, 419.72 - 178.85)
35	m272	(n25, n26)	100 %	(440.02, 282.48)	(645.03 - 317.33, 394.78 - 205.80)
36	m247	(n54, n55)	89 %	(423.82, 194.93)	(537.54 - 321.19, 295.40 - 131.38)
		(n66, n67)	11 %		
37	m550	(n183, n184)	100 %	(402.76, 150.48)	(606.92 - 263.68, 253.39 - 87.56)
38	m41	(n143, n144)	32 %	(401.49 , 292.86)	(491.68 - 329.52, 384.50 - 210.74)
		(n141, n151)	68 %		
39	m71	(n142, n143)	100 %	(394.41, 315.27)	(483.9 - 314.88, 405.16 - 235.75)
40	m646	(n239, n240)	100 %	(381.51, 274.13)	(505.97 - 281.41, 376.26 - 195.07)
41	m619	(n239, n240)	100 %	(381.51, 274.13)	(505.97 - 281.41, 376.26 - 195.07)
42	m44	(n143, n144)	50 %	(378.21, 261.97)	(467.66 - 307.59, 338.86 - 192.30)
		(n141, n151)	50 %		
43	m42	(n143, n144)	45 %	(372.12 , 253.89)	(460.37 - 301.44, 329.70 - 184.75)
		(n141, n151)	55 %		
44	m117	(n36, n40)	63 %	(371.07, 123.28)	(452.61 - 294.37, 167.32 - 86.79)
		(n50, n51)	37 %		
45	m401	(n68, n71)	100 %	(365.38, 206.57)	(482.37 - 277.32, 311.30 - 124.05)
46	m26	(n147, n148)	88 %	(361.39, 326.84)	(430.74 - 299.20, 397.26 - 265.15)
		(n136, n180)	12 %		

Continued

HGT ID	Node in Gene Tree	Recipient Branch in Species Tree (older bound, younger bound)	Recipient Weights	Weighted Mean Ages (Ma) (older bound, younger bound)	95% CI (Ma) (older bound, younger bound)
47	m81	(n153, n157)	100 %	(323.05, 280.01)	(408.56 - 252.13, 362.58 - 209.57)
48	m338	(n13, n14) (n20, n21)	32 % 68 %	(317.05, 186.08)	(406.19 - 233.81, 261.19 - 125.20)
49	m166	(n143, n145)	100 %	(315.27, 225.34)	(405.16 - 235.75, 333.62 - 141.45)
50	m345	(n18, n19)	100 %	(302.11, 40.09)	(413.53 - 214.94, 74.22 - 20.40)
51	m548	(n190, n191)	100 %	(286.33, 234.65)	(380.13 - 219.53, 305.71 - 179.41)
52	m304	(n157, n158)	100 %	(280.01, 224.98)	(362.58 - 209.57, 306.73 - 164.28)
53	m10	(n157, n158)	100 %	(280.01, 224.98)	(362.58 - 209.57, 306.73 - 164.28)
54	m164	(n157, n158)	100 %	(280.01, 224.98)	(362.58 - 209.57, 306.73 - 164.28)
55	m8	(n157, n158)	100 %	(280.01, 224.98)	(362.58 - 209.57, 306.73 - 164.28)
56	m85	(n154, n155)	100 %	(253.41, 215.05)	(336.94 - 183.83, 300.24 - 149.42)
57	m370	(n154, n155)	100 %	(253.41, 215.05)	(336.95 - 183.83, 300.24 -149.42)
58	m371	(n154, n155)	100 %	(253.41, 215.05)	(336.95 - 183.83, 300.24 -149.42)
59	m171	(n154, n155)	100 %	(253.41, 215.05)	(336.95 - 183.83, 300.24 - 149.42)
60	m240	(n63, n64)	100 %	(236.98, 122.65)	(335.45 - 161.76, 207.89 - 74.34)
61	m522	(n191, n192)	100 %	(234.65, 187.77)	(305.71 - 179.41, 247.87 - 142.50)
62	m124	(n50, n51)	100 %	(220.84, 159.53)	(302.13 - 159.01, 226.52 - 106.21)
63	m13	(n155, n156)	100 %	(215.05, 120.26)	(300.24 - 149.42, 190.65 - 72.60)
64	m12	(n155, n156)	100 %	(215.05, 120.26)	(300.24 - 149.42, 190.65 - 72.60)
65	m316	(n27, n30)	100 %	(210.31, 94.52)	(297.65 - 148.08, 170.15 - 47.31)
66	m309	(n27, n28)	100 %	(210.31, 145.90)	(297.65 - 148.08, 216.72 - 98.91)
67	m49	(n147, n150)	100 %	(210.09, 146.05)	(273.32 - 156.01, 204.35 - 97.78)
68	m417	(n51, n52)	100 %	(159.5, 102.61)	(226.52 - 106.21, 161.07 - 61.26)

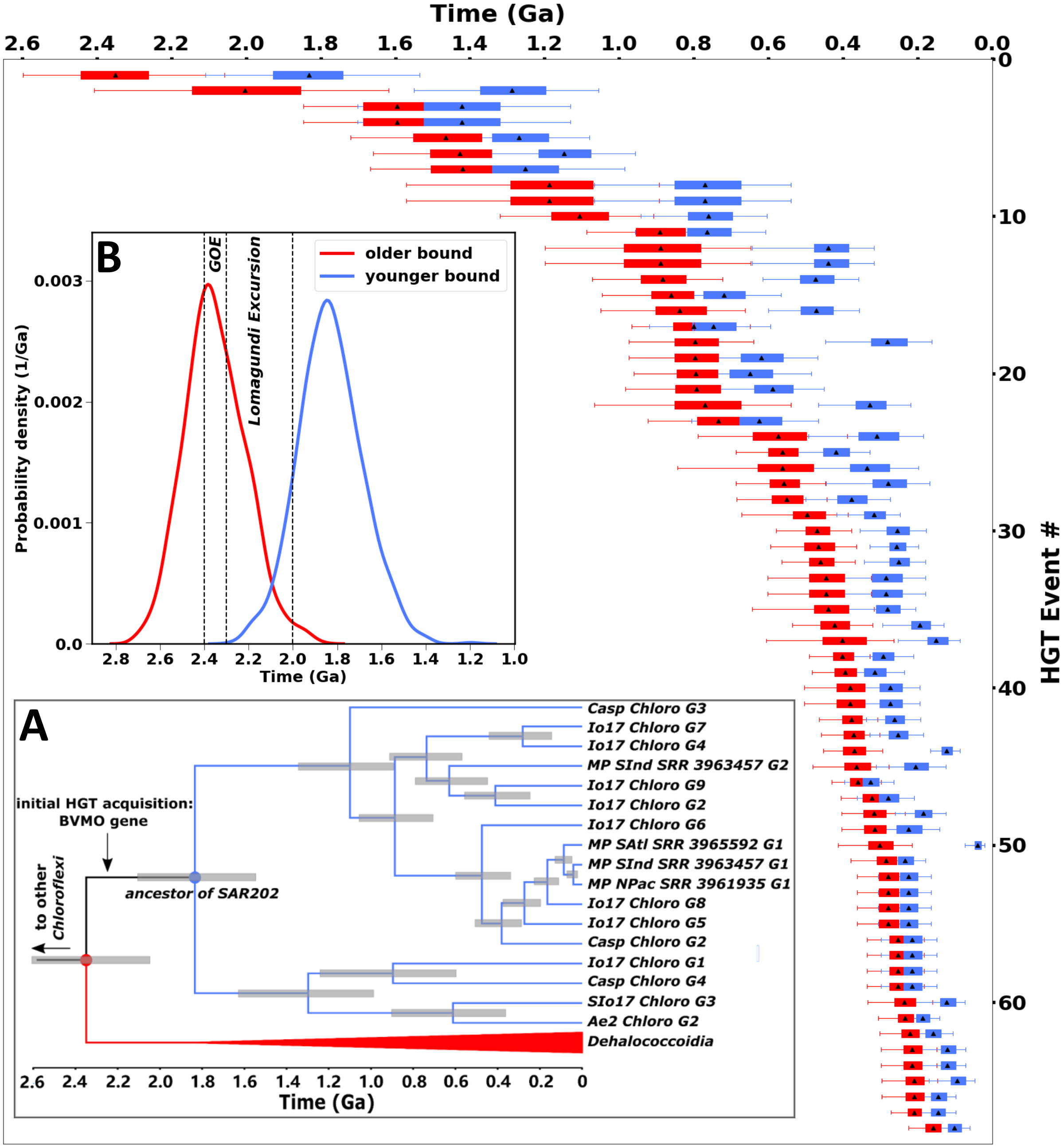


Figure 3-7: The weighted means and 95% CIs of the older and younger time bounds for 68 inferred HGT events (main figure), a chronogram of SAR202 lineages within Chloroflexi (Inset A), and the date distributions of the older and younger age bounds for the initial HGT event into SAR202 (Inset B). The main figure shows a graphic summary for the data in *Appendix SI*, Table S5. Red and blue boxes represents older and younger time bounds, respectively. In these box-plots, black triangles represent means, filled rectangles represent 50% confidence intervals, and horizontal lines represent 95% confidence intervals. Inset A shows a subtree of calibrated chronogram showing SAR202 (blue) and related Dehalococcoidia group (red). The initial HGT acquisition occurred on the branch between stem SAR202 node (red filled circle) and crown node SAR202 (blue filled circle). Grey horizontal bars on the nodes indicate 95% CIs. Inset B Posterior date intervals of the stem (red) and crown (blue) SAR202 nodes shown in the Inset A. Stem and crown date intervals correspond to the distributions of older (red) and younger (blue) time bounds for the initial HGT acquisition shown in the Inset A, and also correspond to the HGT event #1 shown in the main figure. The mean date of older bound is 2350 Ma (95% CI: 2056 Ma - 2598 Ma), and the mean date of younger bound is 1830 Ma (95% CI: 1535 Ma - 2110 Ma). The time windows of the GOE and the Lomagundi Excursion overlap the distributions of older and younger time bounds.

3.5 Diversification Rates

Inset A of Figure 3-7 indicates that, following the initial HGT acquisition, SAR202 BVMO genes diversified at rates that varied non-uniformly with time. Because we expect this variation to be coupled to Earth's redox state, we investigate it in detail in Figure 3-8, which plots the per-gene diversification rate as a function of time. We seek the per-gene diversification rate r, which in continuous time would be defined by

$$r = \frac{1}{N}\frac{dN}{dt} = \frac{d\log N}{dt}, \tag{3.1}$$

where $N(t)$ is the number of nodes at time t after the first 20% of the trees in the "datedist" file were "burned-in". To compute r in discrete time, we denote the average number of nodes at time t_i by N_i, where the bin width $\Delta t = t_{i+1} - t_i = 100$ Myr. Then the diversification rate r_i at time $t_i + \Delta t/2$ is given by

$$r_i = \frac{1}{\Delta t}\log\frac{N_{i+1}}{N_i}. \tag{3.2}$$

Figure 3-8 shows three distinct bursts of diversification of the SAR202 BVMO genes: during the Neoarchean/Paleoproterozoic (around 2500 Ma), the Middle/Late Mesoproterozoic (around 1200 Ma), and the Late Paleozoic/Early Mesozoic (around 300-200 Ma). The timing of these bursts correlates, respectively, with the GOE; the diversification of eukaryotic algae, especially major red and green algae lineages (Sánchez-Baracaldo et al., 2017); and the Permo-Carboniferous O_2 pulse (Berner, 1999). In addition, a large drop in the diversification rate occurred during the period that includes the Neoproterozoic global glaciations (Hoffman et al., 2017).

To assess the significance of the fluctuations of r_i, we consider a null model in which a gene tree randomly diverges at each time step with a constant probability that equals the average diversification rate $\bar{r}$ in Figure 3-8. On average, the number of nodes in the null model grows exponentially with time, like $e^{\bar{r}t}$. However the fluctuations

$$x_i = \log\frac{N_{i+1}}{N_i} - \bar{r}(t_i + \Delta t/2) \tag{3.3}$$

are random and uncorrelated – i.e., they are white noise. In contrast, the fluctuations of the real data are highly correlated in time and are significantly different than white noise. To quantify this difference, we computed the power spectrum (i.e., the amplitude of the Fourier transform, squared) of x_i and compared it to different quantifies of the null-model power spectra. Our numerical analysis (illustrated in Figure 3-9) shows that the probability that the white noise of the null model would generate the highest peak is $p < 0.01$. Moreover, Figure 3-10 shows that the p-values for a Ljung–Box white noise test (Box et al., 2015) at 25 lags are all greater than 0.5. Thus the fluctuations of the diversification rates in Figure 4 are statistically distinct from the results which would be obtained from a randomly-branching gene tree.

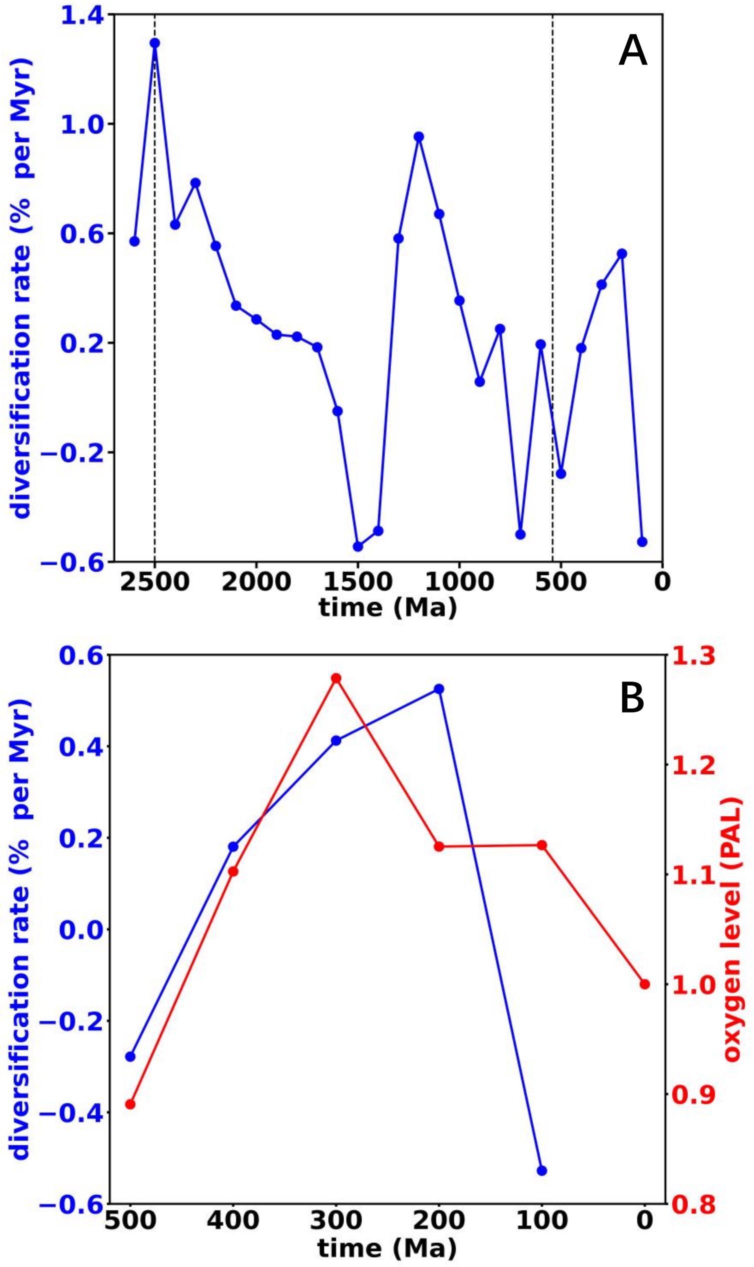


Figure 3-8: (A) The diversification rate of the SAR202 BVMO genes along geologic time. The two vertical dashed lines represent the Archean/Proterozoic boundary and Proterozoic/Phanerozoic boundary, respectively. Three statistically significant bursts likely occurred during: (i) the Neoarchean and Paleoproterozoic (2600 Ma - 2300 Ma), (ii) the Mesoproterozoic (1400 Ma - 1200 Ma), and (iii) the late Paleozoic/early Mesozoic (400 Ma - 200 Ma). These time intervals respectively correspond to (i) the GOE, (ii) the rapid divergence of Eukaryotes, and (iii) the Permo-Carboniferous O_2 pulse. (B) The temporal correlation between the SAR202 BVMO diversification rate (blue) and Earth's atmospheric O_2 level (red) in the Phanerozoic. The red dots are the moving averages of the O_2 levels in (Berner, 1999) with a moving window of 200 Myr, which is the uncertainty range (i.e., 95% confidence interval) of the age distributions in the Phanerozoic on the gene tree chronogram of BVMO homologs (Figure S4).

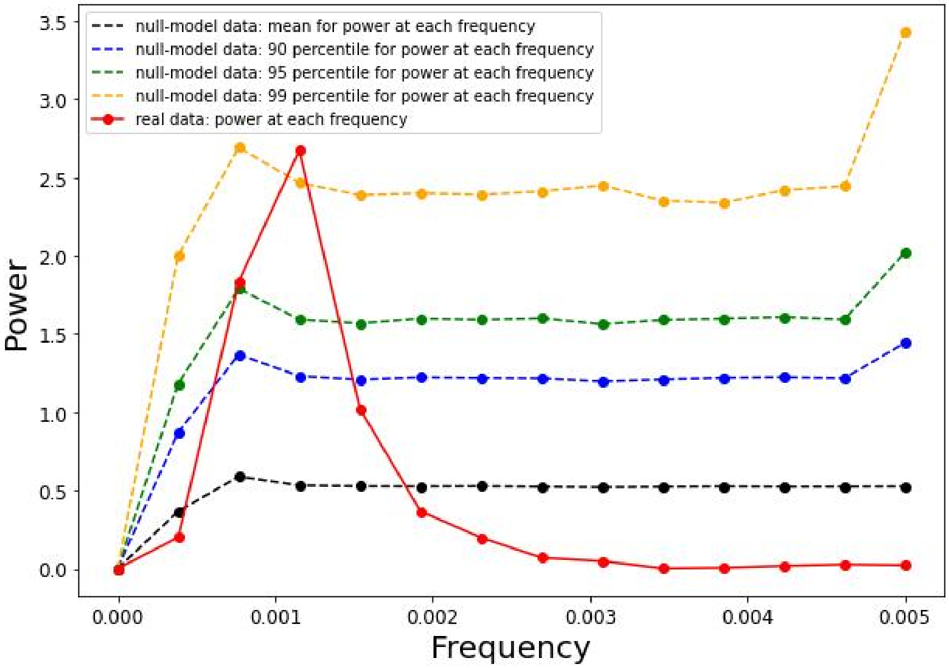


Figure 3-9: A comparison of the power spectra for the real data (i.e., calculated using the posterior age data from the Phylobayes) and the null-model data).

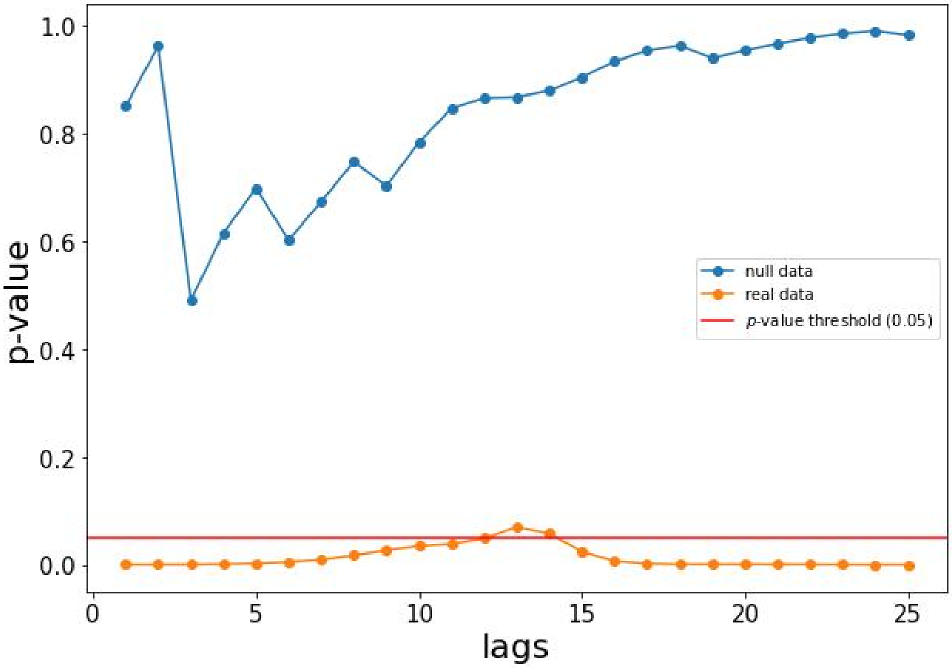


Figure 3-10: A comparison of the p-values for a Ljung–Box white noise test (Box et al., 2015) at 25 lags of the power spectra for the real data (i.e., calculated using the posterior age data from the Phylobayes) and the null-model data).

3.6 Discussion and Conclusion

Our POOM hypothesis suggests a positive feedback for atmospheric O_2 concentrations that derives from the interaction of oxidative metabolic products with sedimentary minerals, and the phylogenetic analyses reveal the temporal correlation between the evolution of POOM-producing microbial metabolisms and the history of Earth's oxygenation. In general, the HGT events (Figure 3-7) and gene diversifications [Figure 3-8-(A)] are shown to positively correlate with the accumulation of O_2 in Earth's atmosphere, as more environments and niches would have sufficient O_2 to maintain oxygen-dependent microbial metabolisms. However, the activity of BVMO genes likely leads to a dichotomous fate of POOM depending on the atmospheric O_2 concentrations. In times when atmospheric O_2 was low and limiting, more POOM was produced as an end product and buried [i.e., $k_{12} > k^\star$ in (2.8)] with the expansion of BVMO genes; in other words, the positive feedback for Earth's oxygenation [the dashed blue curve in Figure 2-5-(B)] was engaged. In contrast, in the well-oxygenated modern environment, the metabolic products of BVMO are more likely to be rapidly, further oxidized before burial [i.e., $k_{12} < k^\star$ in (2.8)], replacing the positive feedback by a negative feedback [the red curve in Figure 2-5-(B)].

Before the GOE, the O_2 concentration in Earth's atmosphere was extremely low and possibly at a steady state. Initially low and localized O_2 production during the Archean likely have instigated the emergence of aerobic metabolic pathways (Taverne et al., 2020). These evolutionary innovations may include the oxidative metabolisms for the production of POOM, such as BVMOs [Figure 3-8-(A)]. Jinich *et al.* (Jinich et al., 2020) have demonstrated that the increase in the redox potential (e.g., the rise of O_2 levels) in an environment promotes the accumulation of microbial metabolites that are rich in carboxyls and hydroxyls. Furthermore, the majority of iron on Earth's surface existed in a reduced form during this time, either as dissolved Fe(II) or within iron(II)-containing minerals such as pyrite (FeS), and the availability of iron(III) oxides was limited (Anbar, 2008). The early rise of atmospheric O_2 enhanced the transformation of iron from ferrous to ferric states on Earth's surface, potentially promoting the physical protection of organic matter through increasing mineral surface charge density (Keil and Mayer, 2014) or enhancing the co-precipitation and chelation of organic matter with Fe(III) oxides (Lalonde et al., 2012). As a result, POOM formed by oxidative metabolisms was strongly protected by the newly accumulated iron(III) minerals. This enhanced organic burial in sediments and led to the elevation of O_2 level in the atmosphere, eventually leading to the GOE and the following Lomagundi Excursion Event. Earth's oxygen cycle then reached a new steady state that lasted for most of the Proterozoic.

Hundreds of millions of years later, a new chapter in the evolutionary history of life began during Late Mesoproterozoic to the Early Neoproterozoic. While atmospheric O_2 levels remained low, the diversification of eukaryotic algae, such as red and green algae, was a key biological feature (Sánchez-Baracaldo et al., 2017). The resistant biopolymers in the cell walls and cysts of these organisms (Knoll, 2014) likely enhanced the portion of recalcitrant organic carbon delivered to sediments and could have instigated a burst in the transfer (Figure 3-7) and diversification [Figure 3-8-(A)] of oxygenases at this time, increasing the production of POOM. However, it is likely that the abundance of iron(III) oxides and other minerals on Earth's surface remained static during this "Boring Billion" (Slotznick et al., 2018). Due to the lack of effective mineral protection, the increase in POOM production was probably not able to initiate a positive feedback and trigger a global oxygenation event. Nevertheless, the environmental stasis in this "Boring Billion" may be a reflection of scarce research and scant geochemical evidence (Zhang et al., 2016). The possibility of local oxygenation event(s) (Zhang et al., 2018) during this geologic time remains to be further investigated.

With the advent of the Neoproterozoic Snowball Earth, anoxic conditions and low primary productivity likely prevailed beneath extensive ice cover (Hoffman et al., 2017), leading the evolution of BVMO genes to stagnate. Consequently, the SAR202 BVMO diversification rate significantly dropped during this time period [Figure 3-8-(A)]. The situation changed when the Late Neoproterozoic arrived. This trend again reversed during the late Cryogenian and Ediacaran following recovery from these glacial periods [Figure 3-8-(A)]. The emergence of new biota, such as mosses and fungi (Heckman et al., 2001), fundamentally changed terrestrial weathering mechanisms (Hazen et al., 2013). This promoted the formation of clay minerals the formation of clay minerals (Hazen et al., 2013), which could enhance the long-term persistence of organic matter (Blattmann et al., 2019). In the meantime, the rapid evolution of multicellular life enhanced the formation of biominerals (Hazen et al., 2013), which could protect organic matter via occlusion (Arnarson and Keil, 2007b). Figure 3-8-(B) suggests that SAR202 BVMO genes underwent another burst of diversification in the Permo-Carboniferous, during which a large amount of organic carbon was buried and the atmospheric O_2 was extraordinarily high (Berner, 1999). The BVMO gene history is but one example that may be indicative of broader evolutionary trends in oxidative microbial metabolisms across these planetary transitions.

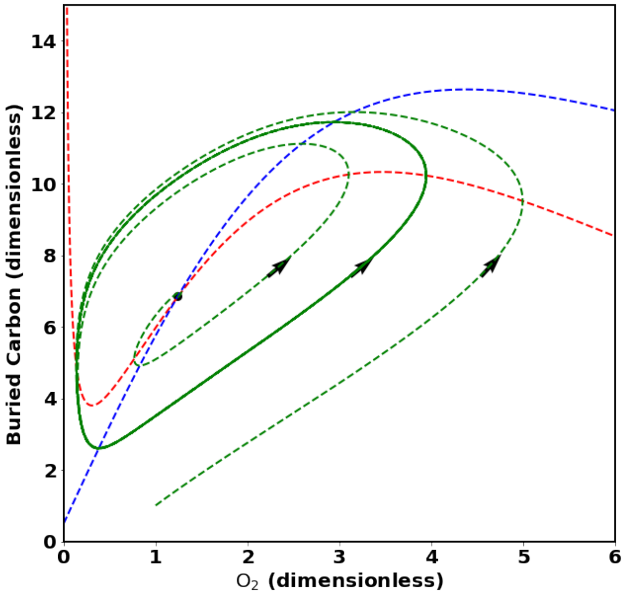
In summary, our POOM hypothesis provides a mechanism for a positive feedback deriving from the interactions of oxidative metabolic products with sedimentary minerals. The temporal consistency of the evolution of an oxidative enzyme family, BVMOs, with the evolution of Earth's atmospheric O_2 supports this hypothesis. Counter-intuitively, the expansion of oxidative metabolisms coupled to the evolution of sedimentary minerals created an autocatalytic dynamics in which the initial advent of oxygen-consuming processes led to oxygen accumulation. The complexity of biogeochemical cycles suggests that similar synergies may help explain other aspects of the episodic character of Earth's biogeochemical evolution.

Chapter 4

Dynamics

Essentially, all models are wrong, but some are useful.

— George Box and Norman Draper, *Empirical Model-Building and Response Surfaces*[*]



Abstract

Chapter 2 has suggested a positive feedback deriving from the expansion of oxidative metabolisms and Chapter 3 has supported it using phylogenomic methods. The dynamics of oxygen and carbon cycles originated from such a positive feedback, however, remains not well understood. Here I construct a mathematical model for Earth's oxygen and carbon cycles and analyze its dynamics, including stability, periodic nonlinear oscillations, and excitability. This model interprets the oxygenation events as dynamical bifurcations, the fluctuations in the oxygen and carbon cycles during oxygenation events as the limit cycles accompanying the unstable transition of Earth's atmospheric O_2 levels, and the large excursions of carbon isotope occurred around the oxygenation events as the oxygen-carbon system's excitation near the bifurcation points.

---

[*] George Box and Norman Draper, *Empirical Model-Building and Response Surfaces.* John Wiley & Sons, New York, NY (1987).

4.1 A Dynamical System of Earth's Carbon and Oxygen Cycles

In this section, we construct a mathematical model to explore the unstable evolution of Earth's oxygen and carbon cycles.

Basic Assumptions

We first make two basic assumptions, which constrain the temporal and spatial scales for the carbon-oxygen system that our model focuses on.
(a) *a well-mixed open system*
We consider a well-mixed open system for the planetary-scale oxygen and carbon cycles on Earth's surface. The words "well mixed" in our context means that we do not have to consider variations in space, which is related to, for example, a "well mixed beaker" where only one variable – time t – is required for the concentration of solution $c(t)$ in the beaker rather than $c(x,t)$, where x is another variable. This assumption implies that the model addresses the biogeochemical cycles with a timescale that is longer than the transfer of carbon between the ocean-atmosphere system and the rock reservoir ($\sim$ 1,000,000 yr) (Berner, 1990, 2004).
(b) *a heat bath analogy*
Since the carbon storage in Earth's mantle and core is much larger than that on its surface (Dasgupta and Hirschmann, 2010; Fischer et al., 2020), we assume that the rate at which reductants or oxidants leave or enter the Earth's mantle and core does not appreciably vary with changes in their mass in the deep Earth. This is analogous to the heat exchange between a small system and a large heat bath. Thus, our model only considers the dynamics of carbon and oxygen cycles on Earth's surface, which is not influenced by the underlying effectively infinite mass of carbon.

Entering and Exiting Fluxes of Organic Carbon in Earth's Crust

We express the flux of buried organic carbon, J_{in}, can be expressed as:

$$J_{\text{in}} = J_p \, B, \tag{4.1}$$

where J_p is the flux of organic carbon deposited on continental shelves and $0 \leq B \leq 1$ is burial efficiency. Previous studies have suggested that the burial of organic matter is not controlled by productivity (Keil et al., 1994; Mayer, 1994; Ransom et al., 1998). In our model, we consider the changes in J_{in} to be due to the variation in the incomplete-degradation fraction of deposited organic carbon (i.e., B) while J_p is maintained as a constant. In the modern environment, $J_p \approx 1.42 \times 10^{13}$mole/yr. Our mathematical model based on the POOM hypothesis predicts the existence of a positive feedback in the ancient oxygen-deficient environment – that is, burial efficiency increases with oxygen concentration (i.e., oxygen-exposure time). Considering the existence of a positive feedback in the ancient O_2-deficient environment, we write the function of burial efficiency as

$$B(x) = \frac{a + bx}{c + x^2}, \tag{4.2}$$

where x is the concentration of O_2 in Earth's atmosphere, and a, b, and c are three parameters with non-negative values. Parameter $0 \leq a/c \leq 1$ represents the burial efficiency of organic carbon in the absence of O_2. Figure 4-1 shows B as a function of x using some appropriate values for the parameters (refer to Table 4.1).

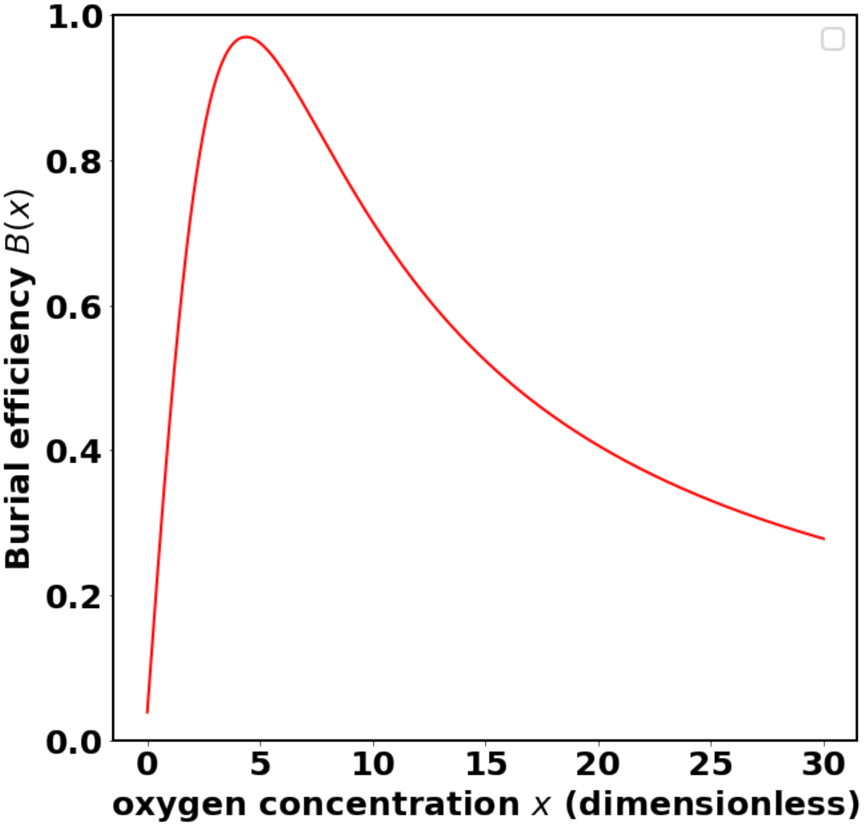


Figure 4-1: Burial efficiency $B(x)$ as a function of dimensionless oxygen concentration x. Refer to Section 4.2 for the definitions of dimensionless variable x and parameters a. b and c. In this plot, the values of the dimensionless a. b and c are taken as 1, 8.5, and 20 (refer to the discussion in Section 4.2).

The maximum value of burial efficiency

$$B_{max} = \frac{1}{2c}\left(a + \sqrt{a^2 + b^2 c}\right) > 0 \tag{4.3}$$

occurs at

$$x_{max} = \frac{\sqrt{a^2 + b^2 c} - a}{b} > 0. \tag{4.4}$$

Substituting (4.2) to (**??**), we obtain the flux of buried organic carbon as a function of O_2 concentration:

$$J_{\text{in}}(x) = J_p \frac{a + bx}{c + x^2}. \tag{4.5}$$

We denote the concentration of organic carbon sequestered as sedimentary rocks in Earth's crust by y. On the tectonic conveyor, which has been established by 2.8 - 2.6 Ga (Kearey et al., 2009) or might be even much older (Brenner et al., 2020), buried organic carbon will be re-exposed to O_2 in continental rifts (Canfield, 2014; Hayes and Waldbauer, 2006). The scale of characteristic time between organic-matter immobilization and the potential for oxidation (that is, the time duration for the transfer of carbon between the ocean-atmosphere system and rock reservoir), denoted here by t_y, is millions of years (Berner, 1990, 2004). The flux of organic carbon recycled back to Earth's surface, J_{out}, then is

$$J_{\text{out}} = \frac{y}{t_y}. \tag{4.6}$$

Source and Sink Fluxes of O_2 in Earth's Atmosphere

The known sources of atmospheric O_2 includes hydrogen escape, pyrite burial and organic burial (Catling and Claire, 2005; Holland, 2006). Hydrogen escape was the dominant mechanism for O_2 production on Earth before the advent of oxygenic photosynthesis (Catling and Claire, 2005), and pyrite burial is a geological process in the sulfur-oxygen system. Since we explore Earth's oxygenation after the emergence of oxygenic photosynthesis, our mathematical model does not include hydrogen escape. Also, because we focus on the carbon-oxygen system, the O_2 production due to pyrite burial is not included in the model. Thus, the source flux of atmospheric O_2 in our model equals the flux of buried organic carbon:

$$F_{\text{source}} = J_{\text{in}}. \tag{4.7}$$

After released into atmosphere, a portion of O_2 will be consumed by reducing gases from volcanoes and metamorphic processes or by recycled organic carbon via oxidative weathering (Canfield, 2014; Catling and Claire, 2005; Lyons et al., 2014). The kinetics of photochemical oxidation of the reducing compounds by O_2 are very fast (Kasting, 1987), thus we assume that the rate of O_2 consumption by the reducing gases from volcanoes and metamorphic processes only depends on the flux of reducing gas and O_2 concentration in Earth's atmosphere. We denote the value for flux of reducing gases as v; the estimated value of present v is $(5.7 \pm 1.2) \times 10^{12}$ mole/yr (Catling and Kasting, 2017). However, v depends on the redox state of Earth's mantle and has been suggested to decrease with geologic time (Aulbach and Stagno, 2016; Kadoya et al., 2020a,b; Nicklas et al., 2019). On the other hand, the rate of O_2 consumption by oxidative weathering can be no faster than the flux of recycled organic carbon potential for oxidative J_{out}. Thus, the maximum rate of O_2 consumption (R) is then expressed

$$R = J_{\text{out}} + v. \tag{4.8}$$

The rate of O_2 consumption by oxidative weathering $F_{\text{sink, w}}$ at O_2 concentration x then can be expressed by

$$F_{\text{sink}}(x) = R \times M(x), \tag{4.9}$$

in which the Michaelis-Menten form reads:

$$M(x) = \frac{x}{k + x}, \tag{4.10}$$

where k is the value of x when $M = 1/2$.

Now we obtain the source flux F_{source} and the sink flux F_{source} as functions of atmospheric O_2

concentration. Eqs.(4.5) and (4.7) together gives the source flux of atmospheric O_2:

$$F_{\text{source}}(x) = J_p \frac{a + bx}{c + x^2}. \tag{4.11}$$

Combing (4.6), (4.8), (4.9) and (4.10) gives the sink flux of atmospheric O_2:

$$F_{\text{source}}(x) = \left(v + \frac{y}{t_y}\right) \frac{x}{k + x}. \tag{4.12}$$

Table 4.1: Model Parameter Values

symbol	physical meaning	value or expression
x	amount of Earth's atmospheric O_2	independent variable
y	amount of organic carbon in Earth's crust	independent variable
a	a parameter controlling B in anoxic environment	–
b	a parameter controlling the maximum burial efficiency	–
c	a parameter controlling B in anoxic environment	–
B	burial efficiency of organic carbon	$(a + bx)/(c + x^2)$
t_y	characteristic recycling time of buried organic carbon	$\sim 10^8$ years
v	flux of reducing agents	control parameter
R	maximum consumption rate of O_2 by weathering	y/t_y
M	a Michaelis-Menten form for O_2 sink flux	$x/(k + x)$
k	a parameter in M; aerobic respiration inhibited ($O_2 \sim 0.01$ PAL)	1.36×10^{19} mol
J_p	organic carbon deposited on the continental shelves	1.42×10^{13}mole/yr
J_{in}	flux of organic carbon buried in sediments	J_pB
J_{out}	flux of organic carbon recycled back to Earth's surface	y/t_y
F_{source}	source flux of atmospheric O_2	J_pB
F_{sink}	sink flux of atmospheric O_2	$(v + R) \times M$

4.2 Dynamics of the Oxygen-Carbon System

This section explores the dynamics of the oxygen-carbon system, including its stability, limit cycles, and excitability. Here I hypothesize that, in Earth's ancient O_2-deficient environment, the burial efficiency at $\chi = 0$ is $B(\chi = 0) = \alpha/\gamma = 0.05$ and I take $\alpha = 1$ and $\gamma = 20$. At the extreme value for the maximum burial efficiency – that is, $B_{max} = 1$, Eq. (4.4) gives $\beta = 8.72$. Here I assume that the mineralization of organic matter had never ceased in the ancient environment and take $\beta = 8.5$ so that the maximum burial efficiency does not reach 1. Although the dynamical system [i.e., Eqs. (4.13) and (4.14)] appears simple, its analytic calculations are challenging (as shown in Section 4.2). The numeric continuation software MATCONT (Dhooge et al., 2003) was used to analyze the dynamical system.

Steady States and Stability

The rates of change of O_2 and organic carbon on Earth's surface are expressed by the difference between their source and sink fluxes:

$$\frac{dx}{dt} = F_{\text{source}} - F_{\text{sink}} = J_p \frac{a + bx}{c + x^2} - \left(\frac{y}{t_y} + v \right) \frac{x}{k + x} \tag{4.13}$$

$$\frac{dy}{dt} = J_{\text{in}} - J_{\text{out}} = J_p \frac{a + bx}{c + x^2} - \frac{y}{t_y}. \tag{4.14}$$

The difference between the flux of organic carbon potential for oxidation and the flux of organic carbon consumed by oxidative weathering is:

$$\frac{y}{t_y} - \frac{y}{t_y} \frac{x}{k + x} > 0, \tag{4.15}$$

Using k as a characteristic concentration and t_y as a characteristic time, we define the following quantities:

$$\chi = \frac{x}{k}, \quad \phi = \frac{y}{k}, \quad \tau = \frac{t}{t_y}, \quad \xi = \frac{J_p t_y}{k}, \quad \alpha = \frac{a}{k^2}, \quad \beta = \frac{b}{k}, \quad \gamma = \frac{c}{k^2}, \quad \nu = \frac{v t_y}{k}. \tag{4.16}$$

Substitute these quantities into (4.13) and (4.14) gives the dimensionless equation system:

$$\frac{d\chi}{d\tau} = \xi \frac{\alpha + \beta\chi}{\gamma + \chi^2} - (\phi + \nu) \frac{\chi}{1 + \chi} \tag{4.17}$$

$$\frac{d\phi}{d\tau} = \xi \frac{\alpha + \beta\chi}{\gamma + \chi^2} - \phi. \tag{4.18}$$

The χ-nullcline is:

$$\phi = \xi \frac{\alpha + \beta\chi}{\gamma + \chi^2} \left(1 + \frac{1}{\chi} \right) - \nu. \tag{4.19}$$

The ϕ-nullcline is:

$$\phi = \xi \frac{\alpha + \beta\chi}{\gamma + \chi^2}. \tag{4.20}$$

The details of stability analyses are provided in Appendix B.

As mentioned above, steady states (or fixed points) occur when $\dot{\chi} = \dot{\phi} = 0$. The oxygen-carbon system exhibits a range of steady states. The red curves in Figure 4-2 show how the stability of the steady states of χ (dimensionless amount of O_2) and ϕ (dimensionless amount of organic carbon) change with the parameter ν. In Figure 4-2, the maximum value of $\nu = 8$ is about 3 times larger than the flux of volcanic and metamorphic gases in modern environments. This maximum value for ν is estimated according to the results in (Holland, 2009). When ν is large, both χ and ϕ are at stable steady states with low levels (i.e., right solid segments on the red curves). As ν decreases, the oxygen-carbon system passes through a Hopf bifurcation (i.e., the right green dots in Figure 4-2) and becomes unstable (i.e., middle dashed segments on the red curves in Figure 4-2). Further decline of ν leads to the oxygen-carbon system to undergo another Hopf bifurcation (i.e., the left

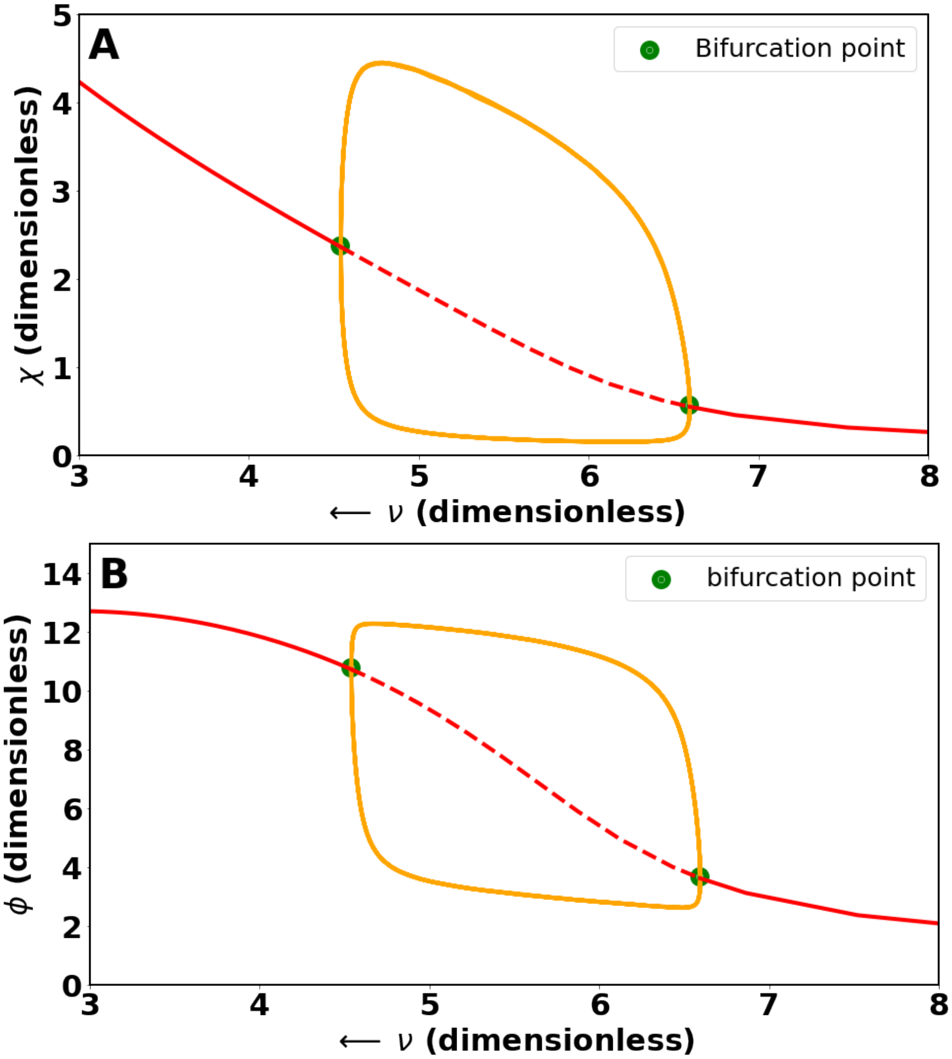


Figure 4-2: Bifurcation diagrams of χ and ϕ with ν as the control parameter. The solid and dashed red curves represent the stable and unstable steady state of χ and ϕ, respectively. The orange solid curves indicate the upper and lower boundaries of the magnitude of the limit cycles. The green dots represent bifurcation points.

green dots in Figure 4-2) and reach new steady states at high levels (i.e., left solid segments on the red curves in Figure 4-2). The flux of reducing agents responsible for O_2 consumption decreased with geologic time (Alcott et al., 2019; Hayes and Waldbauer, 2006; Krissansen-Totton et al., 2015). The flux at 4 billions of years ago was estimated to be $\upsilon = 4.5 \times 10^{13}$ mol/year (Alcott et al., 2019), which equals to a dimensionless value $\nu = (4.5 \times 10^{13}) \times 10^{8}/(1.4 \times 10^{19}) \approx 324$. The present value of υ is estimated to 0.

Limit Cycles

Periodic nonlinear oscillations called limit cycles of χ and ϕ appear when the oxygen-carbon system transitions into the unstable regime. The orange curves in Figure 2 represent the maximum

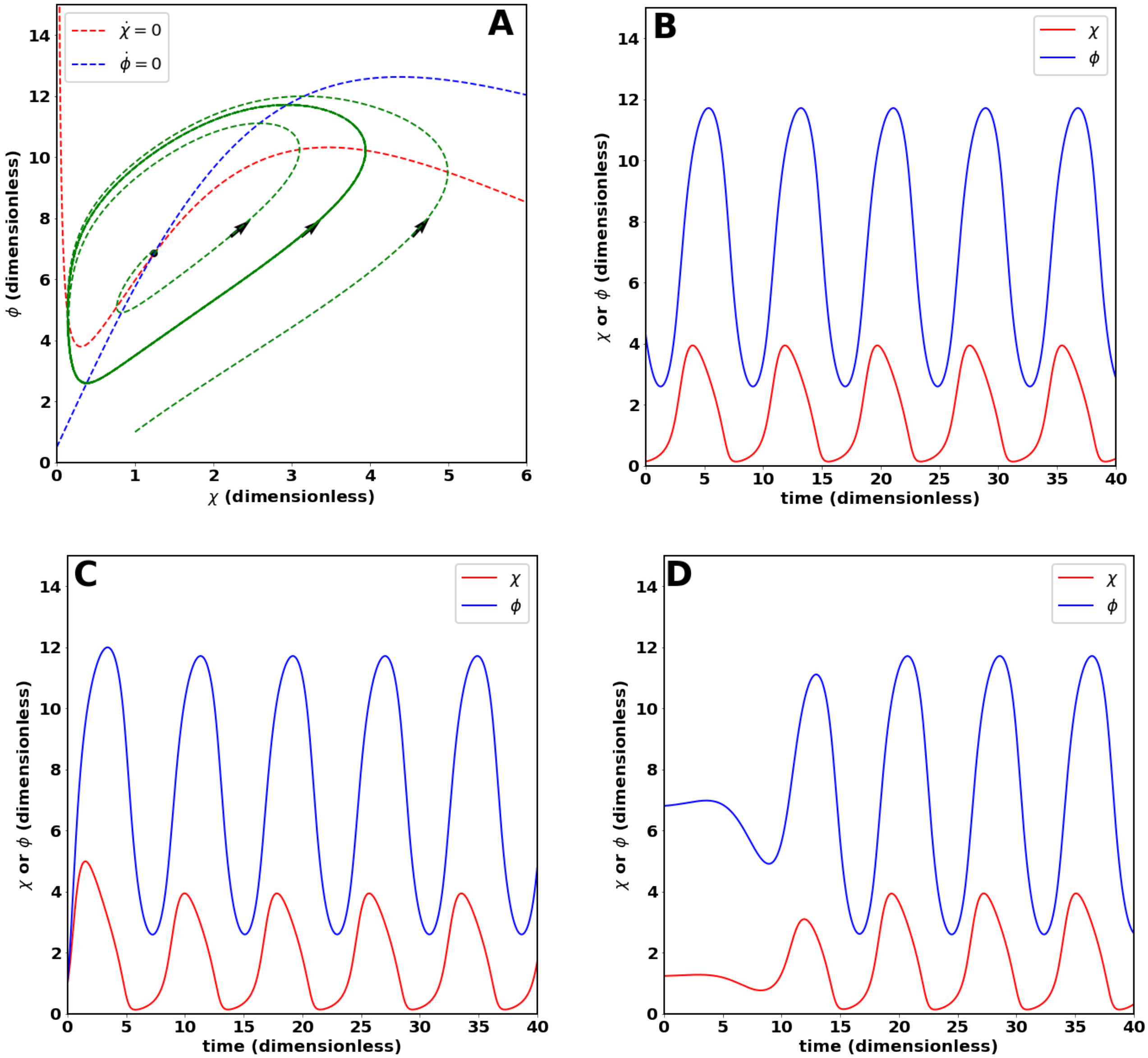


Figure 4-3: (A) A stable limit cycle in the phase plane of dimensionless O_2 level (χ) and dimensionless organic carbon flux (ϕ). The dotted red and blue curves represent nullclines $\dot{\chi} = 0$ and $\dot{\phi} = 0$ when $\upsilon = 5.5$. These two nullclines intersect at an unstable fixed point (the black dot). The solid green curve represents the stable limit cycle. The dotted green curves inside and outside the limit cycle represent two trajectories originating from different initial points that eventually approach the stable limit cycle. Black arrows point to the direction of the trajectories. (B) The time series of χ (red) and ϕ (blue) corresponds to the stable limit cycle in (A). (C) The time series of χ (red) and ϕ (blue) corresponds to the trajectory outside the limit cycle in (A). (D) The time series of χ (red) and ϕ (blue) corresponds to the trajectory inside the limit cycle in (A).

and minimum magnitudes of the limit cycles at each value of the control parameter ν. There are two types of Hopf bifurcations – supercritical and subcritical, in which limit cycles arise in distinct manners. Our analysis shows that the Hopf bifurcations appearing here are supercritical – that is, on one side of which only stable fixed points exist while on the other side of which unstable fixed points coexist with stable limit cycles. It is worth to notice that the amplitude of χ limit cycles (Figure 4-2A) increases as ν decreases while he amplitude of ϕ limit cycles (Figure 4-2B) remains almost constant as ν decreases.

Figure 4-3A depicts a stable limit cycle (i.e., solid green curve) in the phase plane of χ and ϕ. The shape and size of the limit cycle is an intrinsic property of the oxygen-carbon system that does not depend on the initial conditions. For example, the two trajectories (i.e., dotted green curves) shown in Figure 4-3A – one originating inside while the other originating outside the limit cycle – eventually approach this limit cycle. Figure 4-3B demonstrates the time series of χ and ϕ corresponding to the limit cycle in Figure 3A. As typical characteristics of the oxygen-carbon system when it loses its stability through the Hopf bifurcation, the periodic oscillations exhibited by these time series are analogous to the fluctuations in the oxygen and carbon cycles during Earth's oxygenation. This will be further discussed in the next section.

Excitability

Now we explore the response of the oxygen-carbon system when perturbations of the stable steady states occur. When $\nu = 8$, the oxygen-carbon system is at a stable steady state. At time $\tau = 0$, a perturbation to this system leads to a little more initial organic burial ϕ and therefore larger initial χ. However, the state of the system must return to the stable fixed point. Figure 4-4A and B show a small excursion above the fixed point when the trajectory returns to the steady state. Next we set a slightly smaller flux of reducing gases, $\nu = 6.8$; in this case, the stable steady state of the oxygen-carbon system is only marginally lower than the previous situation. The initial condition (due to the small perturbation) remains the same. However, Figure 4-4C and D show that when the system relaxes back to the stable steady state, the size and timescale of the trajectory's excursion above the fixed point is much larger than before and comparable to the limit cycle shown in Figure 4-3.

The above phenomena indicate that the oxygen-carbon system is excitable. A dynamical system is called excitable when it has a trajectory that begins in a small neighboring domain of the steady state, leaves this domain, and finally relaxes back to the steady state after a large-amplitude excursion (Izhikevich, 2007; Strogatz, 2018). Excitability occurs near Hopf bifurcations (Izhikevich, 2007). When the oxygen-carbon system shifts to a region close to the bifurcation points, organic carbon burial occurs as analogous to the great excursions of carbon isotope in the sedimentary records; this will be further discussed in the next section. Here we only present the excitability appearing in the right regime of stable steady states in Figure 4-2 (i.e., when $\nu > 6.6$). The oxygen-carbon system exhibits excitability in a similar manner in the left regime of stable steady states in Figure 4-2 (i.e., when $\nu < 4.5$), which is shown in *SI Appendix*.

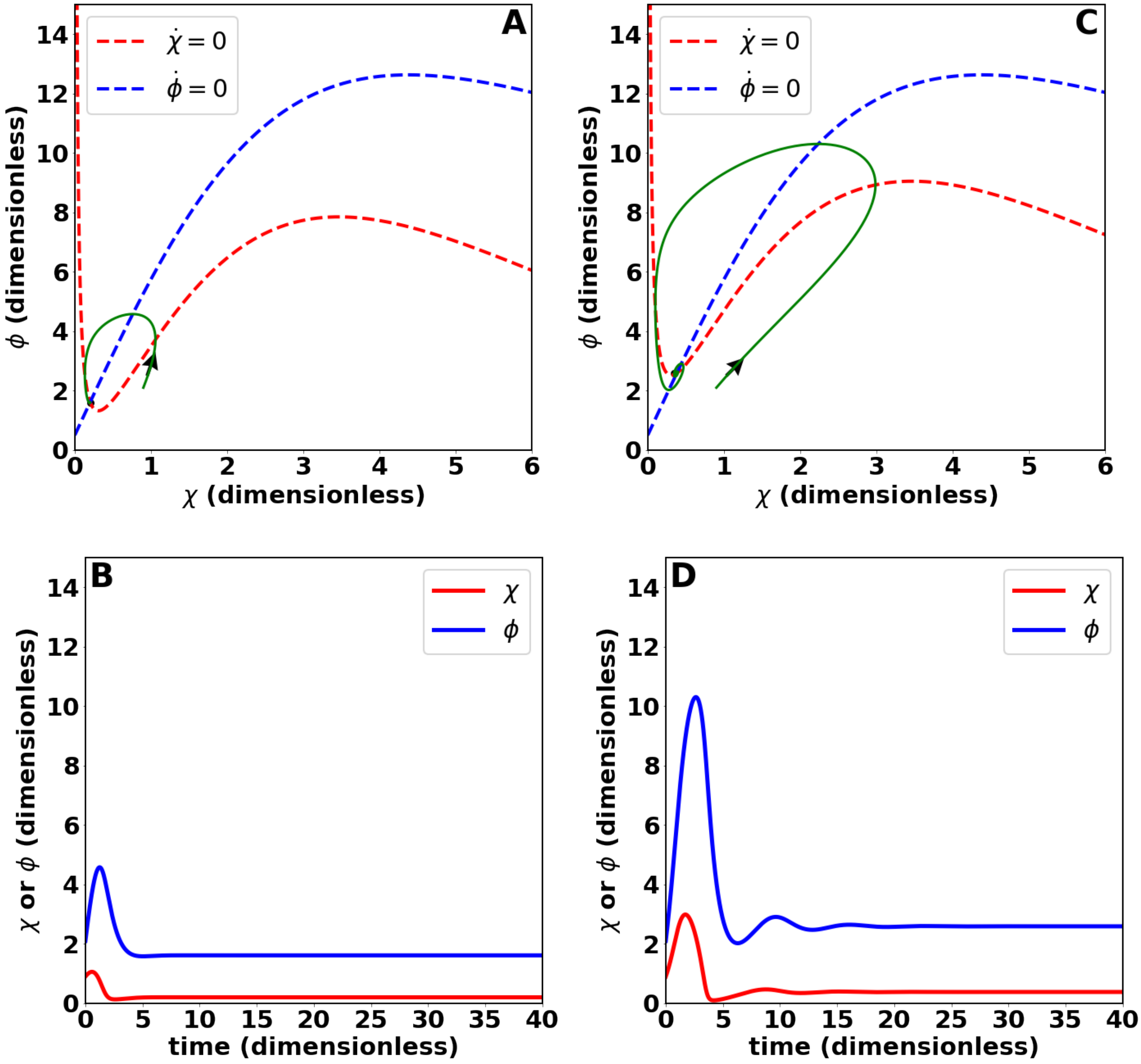


Figure 4-4: Excitability of the oxygen-carbon system at large ν. (A) A trajectory without excitation on the phase plane when $\nu = 8$. (B) Time series of χ and ϕ corresponding to the trajectory in (A). (C) A trajectory with excitation on the phase plane when $\nu = 6.8$ (D) Time series of χ and ϕ corresponding to the trajectory in (C).

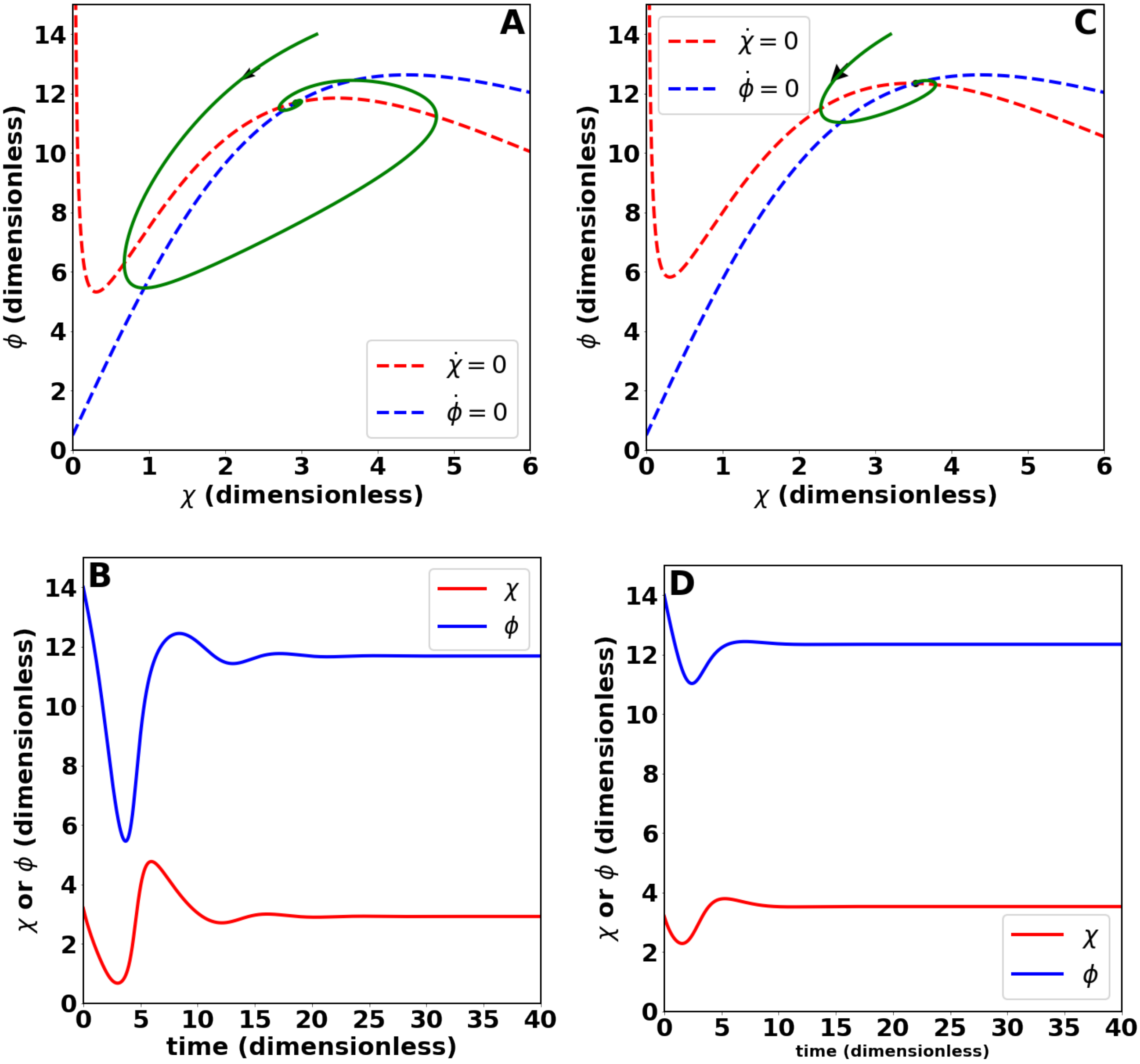


Figure 4-5: Excitability of the oxygen-carbon system at low ν. (A) A trajectory with excitation on the phase plane when $\nu = 4$. (B) Time series of χ and ϕ corresponding to the trajectory in (A). (C) A trajectory without excitation on the phase plane when $\nu = 3.5$. (D) Time series of χ and ϕ corresponding to the trajectory in (C)

4.3 Discussion

Earth's surface was an (almost) O_2-free world before the advent of oxygenic photosynthesis around 2.7-2.8 Ga or even older (Buick, 2008; Farquhar et al., 2011; Kim et al., 2012). Several geochemical signs indicate the existence of a so-called "whiff" of O_2 over the time interval 2.6-2.5 Ga (Anbar et al., 2007; Kaufman et al., 2007; Stüeken et al., 2015), but the atmospheric O_2 level was basically static because there was a great amount of reducing gases and the Earth had to underwent a global redox titration before O_2 could increase (Catling and Kasting, 2017; Holland, 2002; Kasting, 2013). This scenario is manifested by Figure 4-2 (the right regime of fixed points): as the flux of reducing gases declined (ν) declines (Aulbach and Stagno, 2016; Kadoya et al., 2020a,b; Nicklas et al., 2019), the stable steady state of O_2 (χ) very slowly crawls to a higher level.

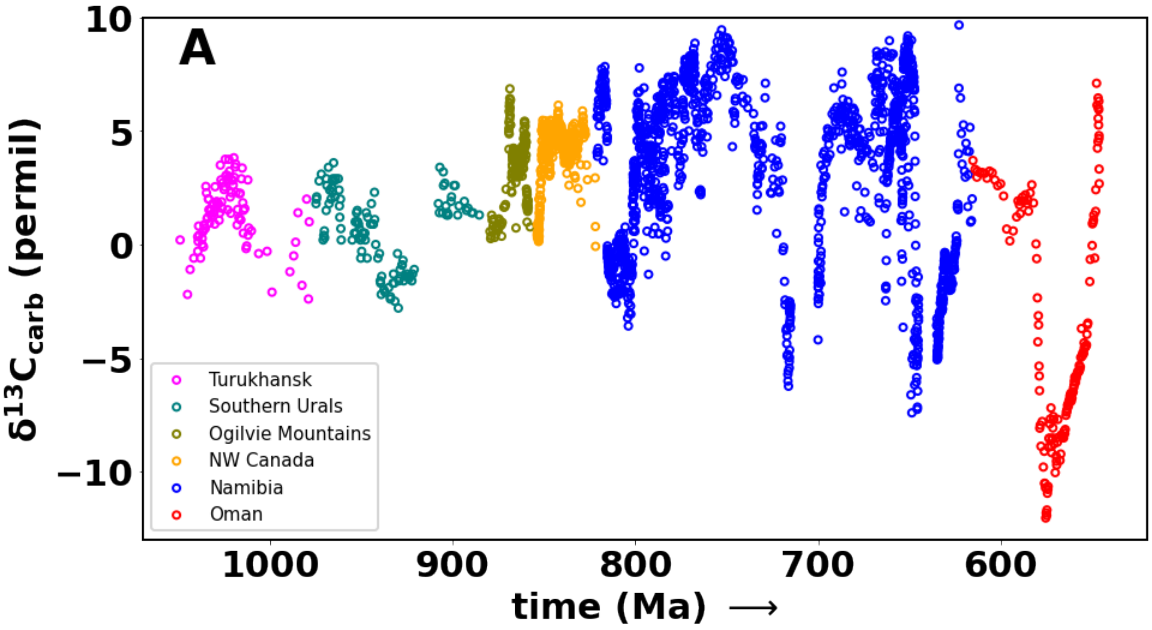


Figure 4-6: Compilation of carbonate $\delta^{13}C_{carb}$ through Neoproterozoic Era. Data source are Turukhansk (Cox et al., 2016; Knoll et al., 1995), Southern Urals (Kuznetsov et al., 2006), Ogilvie Mountains (Cox et al., 2016; Macdonald et al., 2012), Mackenzies Mountains (Halverson, 2006; Rooney et al., 2014), Namibia (Halverson, 2006; Halverson et al., 2005), Oman (Fike et al., 2006), and Morocco (Maloof et al., 2010).

After hundreds of millions of years, a decline of the reducing gases flux triggered the Great Oxidation Event which started in the Late Archean (Aulbach and Stagno, 2016; Kadoya et al., 2020a,b; Kump et al., 2001; Nicklas et al., 2019). The oxygen-carbon system then moved into an excitation region. In this regime, a perturbation of the oxygen-carbon system – e.g., a small increase in organic burial (ϕ) – would drive this system to a neighboring region of its original steady state (Figure 4-4C). The carbon cycle was then excited: the flux of organic carbon buried in the crust rapidly increased to an extraordinarily high level, followed by a drop back to the initial steady state (Figure 4-4 C and D). This excitation phenomenon is analogous to the prominent positive excursion of carbon isotope during the Lomagundi excursion event in 2.2-2.0 Ga (Bachan and Kump, 2015; Lyons et al., 2014), which has been interpreted as a consequence of a dramatic

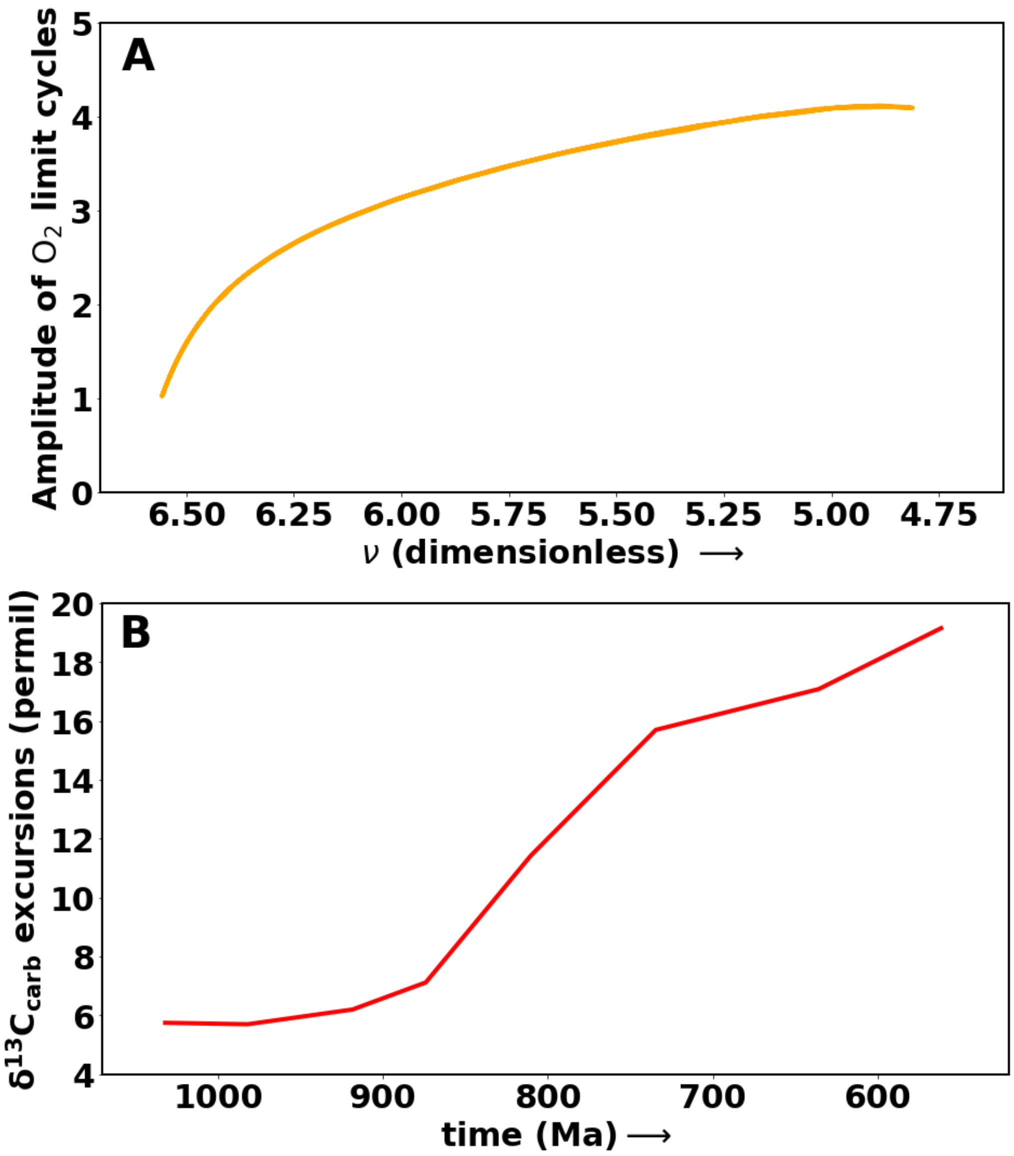


Figure 4-7: A. The amplitude of O_2 limit cycles as a function of ν in the unsatble region [corresponding to the dash red curve in Figure 4-2A]. B. The amplitude of carbon isotope excursions (i.e., the distance between peaks and troughs in Figure 4-6) during the Neoproterozoic Era.

increase in organic burial. The system then moved out from the excitation regime and stayed at a stable state prior to the first Hopf bifurcation (i.e., the right green dots in Figure 4-2). This long-lasting stable steady state is called the “boring billion", during which the oxygen and carbon cycles did not exhibit any significant changes (Holland, 2006).

As the flux of reducing gases further decreased in the late Neoproterozoic (Holland, 2009; Macouin et al., 2015), the oxygen-carbon system passed through the first Hopf bifurcation point (i.e., the right green dots in Figure 4-2). In this unstable regime, the atmospheric O_2 level rose accompanied by fluctuations (Figure 4-2); the periodic nonlinear oscillations of the reducing gases flux (Figure 4-3) are analogous to the prominent excursion events of carbon isotopes during the

Neoproterozoic Era, such as the Garvellach anomaly about 720 Ma [Fairchild et al. (2018); shown in Figure 4-6], the Trezona anomaly about 650 Ga [McKirdy et al. (2001); shown in Figure 4-6], and so on. A mathematical model for Earth's oxygenation by Alcott et al. (2019) also exhibits limit cycles of O_2 levels during the Neoproterozoic oxygenation event. However, different from the limit cycles with almost invariant size (i.e., amplitude) presented in (Alcott et al., 2019), our mathematical model exhibits limit cycles (i.e., the orange curves) whose amplitude increases as the flux of reducing agents declines.

Figure 4-7A shows that the amplitude of limit cycles of O_2 levels increases as ν decreases through the unstable region [the dashed red curve in Figure 4-2A], which provides a theoretical measurement of the atmospheric O_2 levels in this interval. This theoretical prediction not only provides a plausible interpretation for the increasingly deeper $\delta^{13}C_{carb}$ excursions during the Neoproterozoic Era (Figure 4-7B), but also demonstrates a possible, innovative avenue to estimate the changes of the atmospheric O_2 levels based on the information of the carbon isotope data. The increasing amplitude of both the theoretical limit cycles (Figure 4-7A) and the $\delta^{13}C_{carb}$ excursions (Figure 4-7B) are likely a consequence of the increasing size of $\mathbf{O_2}$ reservoir with the geologic time; in other words, the decline of the flux of reductants made more room (in Earth's atmosphere-ocean system) for O_2. As the amplitude of O_2 limit cycle increased, the fluctuation in the amount of remineralized organic carbon became larger, leading to bigger excursions of carbon isotope in sedimentary records. Following this interpretation, we can estimate Earth's atmospheric O_2 levels based on the amplitude of carbon isotope excursions. Figure 4-7B shows that the amplitude of the negative excursions gradually increased during the Neoproterozoic Era and the excursion at the Ediacaran-Cambrian boundary has increased by three times when the oxygen-carbon system leaves the non-steady-state regime. Since the stable O_2 level at the Precambrian-Cambrian Boundary has been estimated to be ∼1 PAL (Canfield, 2014; Lyons et al., 2014), we can estimated that the atmospheric O_2 level at 1 Ga was ∼0.33 PAL. This estimation is consistent with the previous study by Zhang et al. (2016), which has suggested that Earth's atmospheric O_2 level had already been higher than 0.04 PAL at ∼1400 Ma and was sufficient to fuel the evolution of early animals. With a further drop in the reducing gases flux, the oxygen-carbon system passed through the second bifurcation point (i.e., the left green dots in Figure 4-2), moved into a new regime of stable steady states (i.e., the left solid segments on the red curves in Figure 4-2). This is analogous to the new, higher stable state of the atmospheric O_2 levels after the Late Neoproterozoic Oxidation Event.

Chapter 5

Closing Remarks

Your theory is crazy, but it is not crazy enough to be true.

— Niels Bohr*

If the biogeochemical cycles have always been balanced along the geologic time (Simberloff, 2014), how did the Earth evolve from the ancient O_2-deficient world to the modern O_2-rich environment? This question inspires the idea of *autocatalytic* (i.e., unbalanced) evolution of Earth's oxygen cycle introduced in Chapter 1. Generally, autocatalytic phenomena derive from certain positive feedback mechanisms (Prigogine, 1980; Strogatz, 2018). However, observations in the modern environments so far (Hartnett et al., 1998) only show a negative feedback mechanism stabilizing Earth's atmospheric O_2 concentration at the current level. The partially-oxidized organic matter (POOM) hypothesis proposed in Chapter 2 provides a plausible resolution to the enigma of "positive feedback" responsible for the rise of the atmospheric O_2 level in Earth's ancient environment. The relevance of the POOM hypothesis to Earth's autocatalytic oxygenation crucially depends on the existence of POOM-producing oxidative metabolism(s) during the oxygenation events. To test whether such oxidative metabolisms really existed in deep time and to validate the POOM hypothesis, phylogenomic methods were employed in Chapter 3 to reconstruct the evolutionary history of a representative enzyme (i.e., BVMO) that is able to catalyze POOM-producing oxidative metabolisms. The results show a temporal correlation between the evolution of BVMOs and the evolution of Earth's atmospheric O_2, which supports the POOM hypothesis. Based on the POOM hypothesis, a dynamical system model of Earth's oxygen and carbon cycles is provided in Chapter 4. This model interprets the transition of O_2 concentration from a low stable state to a high stable state during the oxygenation events as dynamical bifurcations and the synchronous excursions of carbon isotopes in sedimentary records as characteristic fluctuations associated with the bifurcations. The complexity of biogeochemical cycles suggests that similar synergies may help explain other aspects of the episodic character of Earth's biogeochemical evolution.

The work presented in this thesis provides an innovative theory and evolutionary evidence for the oxygenation of Earth's surface environments. There are several avenues for further investigation of the POOM hypothesis. BVMOs in SAR202 are an important gene family for the production of POOM in the modern marine environments, and have a history that can be traced back across the Phanerozoic and Proterozoic Eons, showing a remarkable degree of phyletic fidelity for a gene family this ancient. For these reasons, the history of BVMOs are especially useful for testing the POOM hypothesis. However, single gene families are inherently limited in the temporal preci-

---

*Spencer Scoular, *First Philosophy: The Theory of Everything.* Universal Publishers, California, USA (2007).

sion that can be attained in using molecular clock dating methods; examining the evolutionary history of other similarly suitable gene families that generate POOM in these environments could therefore potentially provide further support for our hypothesis. Additional laboratory studies of the preservation potential of POOM under conditions analogous to Earth's ancient low-O_2 marine environments could also provide valuable experimental validation. Finally, geochemical investigations of the variation of (mineral-associated) POOM abundance in sedimentary records around the time of Earth's oxygenation events or oceanic anoxic events (Jenkyns, 2010), possibly employing ramped pyrolysis/oxidation (Hemingway et al., 2019), may provide additional opportunities for testing the POOM hypothesis.

Undoubtedly that performing the computational (i.e, phylogenomic), experimental, and field tests suggested above would help us to falsify the POOM hypothesis, but the predictions from a complex mathematical model integrating more variables would give us more chances to validate the hypothesis through the observations on other measurable quantities (Popper, 1959). Although the simple mathematical system presented in thesis provides an elegant interpretation of some dynamics of Earth's oxygen and carbon cycles during the oxygenation events, it is worth to notice that the changes in Earth's atmospheric O_2 concentration did not occur in an isolated system of oxidants and reductants. Some research has suggested that geophysical evolution, such as Earth's rotation rate, surface pressure, continental distribution, ocean salinity and tidal range, might have significantly influenced the trajectory of Earth's oxygenation (Olson et al., 2020). Embedding the equations governing these planetary-scale geophysical variables into the conventional "source-sink" equation system (such as the oxygen-carbon system presented in Chapter 4) for the atmospheric O_2 levels may also display multiple steady states whose stability depends on the variations of these geophysical quantities in addition to the autocatalytic mechanism suggested in this thesis.

It is possible that the positive feedback implied by the POOM hypothesis in Chapter 2 is not the actual mechanisms responsible for Earth's oxygenation events. However, even if this were the case, the value of the oxygen-carbon system model and its dynamics presented in Chapter 4 would not be undermined. This is not because a wrong hypothesis and/or assumption sometimes leads to a correct theory; instead, it is because that the POOM hypothesis and the oxygen-carbon model lie on two different, separated layers of the understanding of Earth's oxygenation events. Some scientists (Wigner, 1950; Yang, 2016) have divided scientific exploration into (at least) four layers: experiments, phenomenological theories, theoretical frameworks, and (pure) mathematics. The positive feedback suggested in Chapter 2 was predicted from a phenomenological theory following the POOM hypothesis, the phylogenomic studies in Chapter 3 then provided a computational experiments to test (and support) this prediction, and Chapter 4 constructed a theoretical framework for Earth's oxygenation. Here I am not stating that the theoretical framework in Chapter 4 is superior to the phenomenological theory in Chapter 2 and the computational experiments in Chapter 3. Instead, I want to claim that the validation of the theoretical framework requires the existence of certain positive feedbacks but does not depend on what the specific positive feedbacks are. If one accepts that the *unstable evolution* is a reasonable view for Earth's oxygenation, then he/she should also agree that the dynamics appearing in the oxygen-carbon system model not only provide plausible interpretations for the abrupt increase in O_2 levels and the accompanying oscillations of the oxygen and carbon cycles during the oxygenation events, but also indicate a possible, innovative avenue of estimating Earth's atmospheric O_2 levels based on the information stored in the carbon isotope records, regardless of the details of the positive feedback mechanisms responsible for Earth's oxygenation. Overall, the conundrum of Earth's oxygenation events investigated in this thesis, other mysterious puzzles of Earth's biogeochemical cycles of oxygen and carbon in deep time (Hayes and Waldbauer, 2006; Lyons et al., 2014; Sessions et al., 2009), and the long-term changing trends of these two coupled cycles in the future all require further exploration. *What's past is prologue*.

All that is transitory
Is but a parable;
The unattainable
Is here attained;
The indescribable
Here is accomplished;
The eternal-feminine
Draws us on high

— Johann Wolfgang von Goethe, *Faust* [†]

[†] Johann Wolfgang von Goethe. *Faust.* Translated by Walter Arndt. W. W. Norton & Company, New York, United States (1898).

Appendix A

Supplementary Information of Chapter 3

Table A.1: List of Ribosomal Proteins

Large Subunit (50 S)	Small Subunit (30 S)
L1	S2
L2	S3
L3	S4
L4	S5
L5	S7
L6	S8
L10	S9
L13	S10
L14	S11
L15	S12
L18	S13
L22	S14
L23	S15
L24	S17
L29	S19

Table A.2: Information of Taxa

Taxon Name	Database	Genome ID
1. *Rickettsia typhi* str. B9991CWPP	PATRIC	1003202.3
2. *Parvularcula bermudensis* HTCC2503	PATRIC	314260.8
3. *Pelagibaca bermudensis* HTCC2601	PATRIC	314265.3
4. *Rhodospirillum centenum* SW	PATRIC	414684.5
5. *Methylobacterium radiotolerans* JCM 2831	PATRIC	426355.14
6. *Caulobacter segnis* ATCC 21756	PATRIC	509190.6
7. *Candidatus Pelagibacter ubique* HTCC1062	PATRIC	335992.9
8. *Pedobacter kyungheensis* KACC 16221	PATRIC	1069985.4
9. *Blattabacterium* sp. (*Cryptocercus punctulatus*) str. Cpu	PATRIC	1075399.3
10. *Formosa* sp. AK20	PATRIC	1137281.3
11. *Mangrovimonas yunxiaonensisstr.* LY01	PATRIC	1197477.3
12. *Cellulophaga geojensis* KL-A	PATRIC	1328323.3
13. *Cardinium endosymbiont* cBtQ1 of Bemisia tabaci	PATRIC	1354314.3
14. *Elizabethkingia meningoseptica* 502	PATRIC	1355388.3
15. *Rikenellaceae bacterium* M3	PATRIC	1433126.3
16. *Flavobacterium aquatile* LMG 4008	PATRIC	1453498.8
17. *Schleiferia thermophila* str. Yellowstone	PATRIC	1453500.3
18. *Chryseobacterium antarcticum*	PATRIC	266748.6
19. *Capnocytophaga canimorsus* str. CcD38	PATRIC	28188.6
20. *Robiginitalea biformata* HTCC2501	PATRIC	313596.4
21. *Porphyromonas canoris* COT-108$_{O}H1224$	PATRIC	36875.4
22. *Kordia algicida* OT-1	PATRIC	391587.3
23. *Epilithonimonas lactis* str. LMG 24401	PATRIC	421072.8
24. *Chryseobacterium luteum* str. DSM 18605	PATRIC	421531.5
25. *Parabacteroides distasonis* ATCC 8503	PATRIC	435591.13
26. *Bacteroides fragilis* $3_{11}2$	PATRIC	457424.5
27. *Spirosoma linguale* DSM 74	PATRIC	504472.7
28. *Nonlabens dokdonensis* DSW-6	PATRIC	592029.3
29. *Marivirga tractuosa* DSM 4126	PATRIC	643867.3
30. *Paludibacter propionicigenes* WB4	PATRIC	694427.4
31. *Paraprevotella clara* YIT 11840	PATRIC	762968.3
32. *Weeksella virosa* DSM 16922	PATRIC	865938.3

Continued

Taxon Name	Database	Genome ID
33. *Ornithobacterium rhinotracheale* DSM 15997	PATRIC	867902.3
34. *Niabella soli* DSM 19437	PATRIC	929713.3
35. *Prevotella nigrescens* ATCC 33563	PATRIC	997352.4
36. *Fluviicola taffensis* DSM 16823	PATRIC	755732.3
37. *Owenweeksia hongkongensis* DSM 17368	PATRIC	926562.3
38. *Blattabacterium* sp. (*Blaberus giganteus*)	PATRIC	1186051.3
39. *Chlorobium tepidum* TLS	PATRIC	194439.7
40. *Chlorobium phaeobacteroides* BS1	PATRIC	331678.5
41. *Chlorobium ferrooxidanDSM* 13031	PATRIC	377431.3
42. *Chlorobium chlorochromatii* CaD3	PATRIC	340177.11
43. *Chlorobium phaeovibrioides* DSM 265	PATRIC	290318.6
44. *Chlorobium luteolum* DSM 273	PATRIC	319225.5
45. *Chlorobium phaeobacteroides* DSM 266	PATRIC	290317.9
46. *Chlorobaculum parvum* NCIB 8327	PATRIC	517417.5
47. *Chlorobium limicola* DSM 245	PATRIC	290315.5
48. *Pelodictyon phaeoclathratiforme* BU-1	PATRIC	324925.5
49. *Chloroherpeton thalassium* ATCC 35110	PATRIC	517418.5
50. *Prosthecochloris aestuarii* DSM 271	PATRIC	290512.6
51. *Nitrolancetus hollandicus* Lb	PATRIC	1129897.3
52. *Dehalococcoides mccartyi* CG1	PATRIC	1432059.3
53. *Thermomicrobium roseum* DSM 5159	PATRIC	309801.4
54. *Chloroflexus aurantiacus* J-10-fl	PATRIC	324602.8
55. *Chloroflexus aggregans* DSM 9485	PATRIC	326427.4
56. *Sphaerobacter thermophilus* DSM 20745	PATRIC	479434.6
57. *Ktedonobacter racemifer* DSM 44963	PATRIC	485913.3
58. *Dehalogenimonas lykanthroporepellens* BL-DC-9	PATRIC	552811.9
59. *Caldilinea aerophila* DSM 14535 = NBRC 104270	PATRIC	926550.5
60. *Anaerolinea thermophila* UNI-1	PATRIC	926569.3
61. *Roseiflexus* sp. RS-1	PATRIC	357808.5
62. *Roseiflexus castenholzii* DSM 13941	PATRIC	383372.6
63. *Prosthecochloris aestuarii* DSM 271	PATRIC	290512.6
64. *Leptolyngbya* sp. PCC 7375	PATRIC	102129.3

Continued

Taxon Name	Database	Genome ID
65. *Nostoc* sp. PCC 7120	PATRIC	103690.1
66. *Stanieria cyanosphaera* PCC 7437	PATRIC	111780.3
67. *Geitlerinema* sp. PCC 7407	PATRIC	1173025.3
68. *Coleofasciculus chthonoplastes* PCC 7420	PATRIC	118168.3
69. *Prochlorococcus marinus* subsp. marinus str. CCMP1375	PATRIC	167539.5
70. *Thermosynechococcus elongatus* BP-1	PATRIC	197221.4
71. *Trichodesmium erythraeum* IMS101	PATRIC	203124.6
72. *Gloeobacter violaceus* PCC 7421	PATRIC	251221.4
73. *Chroococcidiopsis thermalis* PCC 7203	PATRIC	251229.3
74. *Synechococcus elongatus* PCC 6301	PATRIC	269084.6
75. *Cyanobium gracile* PCC 6307	PATRIC	292564.3
76. *Rivularia* sp. PCC 7116	PATRIC	373994.3
77. *Cyanothece* sp. PCC 7822	PATRIC	497965.6
78. *Arthrospira platensis* NIES-39	PATRIC	696747.3
79. *Pseudanabaena* sp. PCC 7367	PATRIC	82654.3
80. *Melioribacter roseus* P3M-2	PATRIC	1191523.3
81. *Ignavibacterium album* JCM 16511	PATRIC	945713.3
82. *Obscuribacter phosphatis* (Mle1_12)	IMG	2556921048
83. *Gastranaerophilaceae* Zag_1 (Zagget bin 1)	IMG	2523533517
84. *Gastranaerophilaceae Zag_111* (Zagget_111_MP)	IMG	2531839741
85. *Gastranaerophilus phascolarctosicola* (Zagget bin 221)	IMG	2523533519
86. *Sericytochromatia*, from Coal Bed ML635J-21	R. Soo (Soo et al., 2017)	CBMW_12
87. *Sericytochromatia*, from Rifle Acetate Amendment	R. Soo (Soo et al., 2017)	CBMW_12
88. *Sericytochromatia*, from Photoreactor Bin 72	R. Soo (Soo et al., 2017)	LSPB_72
89. *Oscillochloris trichoides*	NCBI	GCF_000152145.1_ASM15214v1
90. *Dictyobacter aurantiacus*	NCBI	GCA_003967515.1_ASM396751v1
91. *Thermosporothrix hazakensis*	NCBI	GCA_003253565.1_ASM325356v1
92. *Ktedonobacter* sp. 13_2_20CM_53_11	NCBI	GCA_001914715.1_ASM191471v1
93. *Thermogemmatispora carboxidivorans*	NCBI	GCA_000702505.1_ASM70250v1
94. *Thermogemmatispora onikobensis*	NCBI	GCA_001748285.1_ASM174828v1
95. *Thermogemmatispora tikiterensis*	NCBI	GCA_003268475.1_ASM326847v1
96. *Kouleothrix aurantiaca*	NCBI	GCA_001399705.1_ASM139970v1

Continued

Taxon Name	Database	Genome ID
97. *Candidatus Chloroploca asiatica*	NCBI	GCA_002532075.1_ASM253207v1
98. *Candidatus Viridilinea halotolerans*	NCBI	GCA_003934145.1_ASM393414v1
99. *Chloroflexus islandicus*	NCBI	GCA_001650695.1_ASM165069v1
100. *Herpetosiphon geysericola*	NCBI	GCA_001306135.1_ASM130613v1
101. *Herpetosiphon llansteffanense*	NCBI	GCA_003205875.1_ASM320587v1
102. *Thermorudis peleae*	NCBI	GCA_000744775.1_ASM74477v1
103. *Litorilinea aerophila*	NCBI	GCA_002148365.1_ASM214836v1
104. *Ardenticatena maritima*	NCBI	GCA_001293545.1_ASM129354v1
105. *Ardenticatenia bacterium*	NCBI	GCA_003130875.1_ASM313087v1
106. *Thermoflexia bacterium*	NCBI	GCA_003694475.1_ASM369447v1
107. *Anaerolineae bacterium* SG8_19	NCBI	GCA_001303105.1_ASM130310v1
108. *Candidatus Promineofilum breve*	NCBI	GCA_900066015.1_Cfx-K
109. *Anaerolineaebacterium* SM23_63	NCBI	GCA_001303965.1_ASM130396v1
110. *Brevefilum fermentans*	NCBI	GCA_900184705.1_CAMBI-1
111. *Anaerolineae bacterium* UTCFX2	NCBI	GCA_002050125.1_ASM205012v1
112. *Flexilinea flocculi*	NCBI	GCA_001192795.1_ASM119279v1
113. *Levilinea saccharolytica*	NCBI	GCA_001306035.1_ASM130603v1
114. *Ornatilinea apprima*	NCBI	GCA_001306115.1_ASM130611v1
115. *Longilinea arvoryzae*	NCBI	GCA_001050235.2_ASM105023v2
116. *Leptolinea tardivitalis*	NCBI	GCA_001050275.1_ASM105027v1
117. *Thermanaerothrix daxensis*	NCBI	GCA_001306145.1_ASM130614v1
118. *Bellilinea caldifistulae*	NCBI	GCA_001306055.1_ASM130605v1
119. *Pelolinea submarina*	NCBI	GCA_003966975.1_$ASM396697v1$
120. *Anaerolinea* sp. 4484_236	NCBI	GCA_002085065.1_ASM208506v1
121. *Anaerolineae bacterium* CG2_30_58_95	NCBI	GCA_001872165.1_ASM187216v1
122. *SAR202 cluster bacterium* Io17-Chloro-G1	NCBI	GCA_002816375.1_ ASM281637v1
123. *SAR202 cluster bacterium* Io17-Chloro-G2	NCBI	GCA_002816875.1_ASM281687v1
124. *SAR202 cluster bacterium* Io17-Chloro-G3	NCBI	GCA_002816985.1_ ASM281698v1
125. *SAR202 cluster bacterium* Io17-Chloro-G4	NCBI	GCA_002816455.1_ ASM281645v1
126. *SAR202 cluster bacterium* Io17-Chloro-G5	NCBI	GCA_002816355.1_ ASM281635v1
127. *SAR202 cluster bacterium* Io17-Chloro-G6	NCBI	GCA_002816585.1_ ASM281658v1
128. *SAR202 cluster bacterium* Io17-Chloro-G7	NCBI	GCA_002817055.1_ ASM281705v1

Continued

Taxon Name	Database	Genome ID
129. *SAR202 cluster bacterium* Io17-Chloro-G8	NCBI	GCA_002817135.1_ ASM281713v1
130. *SAR202 cluster bacterium* Io17-Chloro-G9	NCBI	GCA_002816675.1_ ASM281667v1
131. *SAR202 cluster bacterium* Casp-Chloro-G2	NCBI	GCA_002816575.1_ ASM281657v1
132. *SAR202 cluster bacterium* Casp-Chloro-G3	NCBI	GCA_002816715.1_ ASM281671v1
133. *SAR202 cluster bacterium* Casp-Chloro-G4	NCBI	GCA_002816705.1_ ASM281670v1
134. *SAR202 cluster bacterium* Ae2-Chloro-G2	NCBI	GCA_002816925.1_ ASM281692v1
135. *SAR202 cluster bacterium* MP-SInd-SRR3963457-G1	NCBI	GCA_002817235.1_ ASM281723v1
136. *SAR202 cluster bacterium* MP-SInd-SRR3963457-G2	NCBI	GCA_002816735.1_ ASM281673v1
137. *SAR202 cluster bacterium* MP-NPac-SRR3961935-G1	NCBI	GCA_002817285.1_ ASM281728v1
138. *SAR202 cluster bacterium* MP-SAtl-SRR3965592-G1	NCBI	GCA_002817125.1_ ASM281712v1
139. *Fischerella* sp. JSC-11	NCBI	GCF_000231365.1_ASM23136v1
140. *Prochlorothrix hollandicaPCC* 9006 = CALU 1027	NCBI	GCF_000332315.1_ ASM33231v1
141. *Wolbachia endosymbiont of Drosophila simulans* wNo	NCBI	GCA_000376585.1_ ASM37658v1
142. *Anaplasma phagocytophilum* str. Webster	NCBI	GCA_000964685.1_ ASM96468v1
143. *Magnetococcus marinus* MC-1	NCBI	GCA_000014865.1_ASM1486v1
144. *Ehrlichia canis* str. Jake	NCBI	GCA_000012565.1_ASM1256v1
145. *Gemmatimonadetes bacterium*	NCBI	GCA_003134685.1_20111000$_P3M$
146. *Candidatus Entotheonella factor*	NCBI	GCA_000522425.1v3
147. *Candidatus Entotheonella* sp. TSY2	NCBI	GCA_000522445.1v3
148. *Candidatus Entotheonella palauensis*	NCBI	GCA_900079095.1_E.gemina.ts
149. *Candidatus Entotheonella serta*	NCBI	GCA_002914425.1_ASM291442v1
150. *Candidatus Rokubacteria bacterium* 13_1_40CM_2_68_8	NCBI	GCA_001919035.1_ASM191903v1
151. *Rhodospirillales bacterium* 35-66-84	NCBI	GCA_002281325.1_ASM228132v1
152. *Reyranella* sp.	NCBI	GCA_004297405.1_ASM429740v1
153. *Reyranella massiliensis*	NCBI	GCA_000312425.1_ASM31242v1
154. *Enhydrobacter aerosaccus*	NCBI	GCA_900167455.1_IMG-taxon_2582581318
155. *Alphaproteobacteria bacterium* 65-37	NCBI	GCA_001897995.1_ASM189799v1
156. *Rhodospirillaceae bacterium* CCH5-H10	NCBI	GCA_001557035.1_ASM155703v1
157. *Rhodospirillales bacterium* 69-11	NCBI	GCA_001898425.1_ASM189842v1
158. *Belnapia moabensis*	NCBI	GCA_000745835.1_ASM74583v1
159. *Belnapia* sp. F-4-1	NCBI	GCA_000802185.1_ASM80218v1
160. *Rhodospirillaceae bacterium* SYSU D60015	NCBI	GCA_003576705.1_ASM357670v1

Continued

Taxon Name	Database	Genome ID
161. *Thermoflexus hugenholtzii*	NCBI	GCA_002563855.1_ASM256385v1
162. *Rhodospirillales bacterium* URHD0017	NCBI	GCA_900110395.1_IMG-taxon_2590828823
163. *Humitalea rosea*	NCBI	GCA_003253705.1_ASM325370v1
164. *Roseomonas stagni*	NCBI	GCA_900114315.1_IMG-taxon_2602042029
165. *Rhodospirillales bacterium* 70-18	NCBI	GCA_001899585.1_ASM189958v1
166. *Paracraurococcus ruber*	NCBI	GCA_004353985.1_ASM435398v1
167. *beta proteobacterium MWH-P2sevCIIIb*	NCBI	GCA_003003055.1_ASM300305v1
168. *Rhodospirillales bacterium* 69-11	NCBI	GCA_001898425.1_ASM189842v1
169. *Rhodospirillaceae bacterium* SYSU D60007	NCBI	GCA_003574685.1_ASM357468v1
170. *Hyphomicrobium* sp. CS1BSMeth3	NCBI	GCA_900117415.1_CS1BSMeth3_assembly_1
171. *Streptomyces* sp. Ag109_O5-10	NCBI	GCA_900105755.1_IMG-taxon_2675902962
172. *Acetobacteraceae bacterium*	NCBI	GCA_004843345.1_ASM484334v1
173. *Dankookia rubra*	NCBI	GCA_004355005.1_ASM435500v1
174. *Paracraurococcus* sp. NE82	NCBI	GCA_004343615.1_ASM434361v1
175. *Streptomyces fulvoviolaceus*	NCBI	GCA_000718165.1_ASM71816v1
176. *Tepidicaulis* sp. EA10	NCBI	GCA_003688365.1_ASM368836v1
177. *Roseococcus* sp. SYP-B2431	NCBI	GCA_004336745.1_ASM433674v1
178. *Streptomyces mirabilis*	NCBI	GCA_000746395.1_ASM74639v1
179. *Streptomyces* sp. SA15	NCBI	GCA_002291145.1_ASM229114v1
180. *Cystobacter fuscus*	NCBI	GCA_002305875.1_ASM230587v1
181. *Streptomyces albus*	NCBI	GCA_000827005.1_ASM82700v1
182. *Mycobacterium* sp. AB57	NCBI	GCA_900157375.1_PRJEB19165
183. *Streptomyces puniciscabiei*	NCBI	GCA_001735805.1_ASM173580v1
184. *Alphaproteobacteria bacterium* HGW-Alphaproteobacteria-3	NCBI	GCA_002842875.1_ASM284287v1
185. *Frankia* sp. EUN1f	NCBI	GCA_000177675.1_ASM17767v1
186. *Parvibaculum lavamentivorans*	NCBI	GCA_000017565.1_ASM1756v1
187. *Mycolicibacterium moriokaense*	NCBI	GCA_002086395.1_ASM208639v1
188. *Acidimicrobiaceae bacterium*	NCBI	GCA_002694625.1_ASM269462v1
189. *Phenylobacterium* sp. LX32	NCBI	GCA_003254475.1_ASM325447v1
190. *Streptomyces* sp. SM1	NCBI	GCA_002910825.1_ASM291082v1
191, *Streptomyces rimosus*	NCBI	GCA_004196335.1_ASM419633v1
192. *Alphaproteobacteria bacterium* HGW-Alphaproteobacteria-11	NCBI	GCA_002841195.1_ASM284119v1

Continued

Taxon Name	Database	Genome ID
193. *Euzebya tangerina*	NCBI	GCA_003074135.1_ASM307413v1
194. *Ilumatobacter fluminis*	NCBI	GCA_004364865.1_ASM436486v1
195. *Rhodosalinus* sp. E84	NCBI	GCA_003298775.1_ASM329877v1
196. *Mycobacterium palustre*	NCBI	GCA_002101785.1_ASM210178v1
197. *Ilumatobacter nonamiensis*	NCBI	GCA_000350145.1_ASM35014v1
198. *Mycobacterium* sp. 1274761.0	NCBI	GCA_001668615.1_SM166861v1
199. *Mycobacterium triplex*	NCBI	GCA_002102415.1_ASM210241v1
200. *Caulobacteraceae bacterium*	NCBI	GCA_003164925.1_20120700_S1D
201. *Sphingomonas mucosissima*	NCBI	GCA_002197665.1_ASM219766v1
202. *Salinispora pacifica*	NCBI	GCA_000514775.1_ASM51477v1
203. *Caulobacter mirabilis*	NCBI	GCA_002749615.1_ASM274961v1
204. *Rubritepida flocculans*	NCBI	GCA_000425365.1_ASM42536v1
205. *Phenylobacterium* sp. SCN 70-31	NCBI	GCA_001724605.1_ASM172460v1
206. *Sphingomonas* sp. 66-10	NCBI	GCA_001897045.1_ASM189704v1
207. *Phenylobacterium* sp. RIFCSPHIGHO2_01_FULL_69_31	NCBI	GCA_001824475.1_ASM182447v1
208. *Mycobacterium* sp. 1100029.7	NCBI	GCA_001665235.1_ASM166523v1
209. *Sphingomonas* sp. 67-36	NCBI	GCA_001897375.1_ASM189737v1
210. *Cryptosporangium aurantiacum*	NCBI	GCA_900143005.1_IMG-taxon_2700988713
211. *Phenylobacterium zucineum*	NCBI	GCA_003243355.1_ASM324335v1
212. *Sphingomonas jeddahensis*	NCBI	GCA_001981525.1_ASM198152v1
213. *Frankia asymbiotica*	NCBI	GCA_001983105.1_ASM198310v1
214. *Mycobacterium arosiense*	NCBI	GCA_002086125.1_ASM208612v1
215. *Mycobacterium mantenii*	NCBI	GCA_002086335.1_ASM208633v1
216. *Mycobacterium marseillense*	NCBI	GCA_002285715.1_ASM228571v1
217. *Streptomyces scabrisporus*	NCBI	GCA_000372745.1_ASM37274v1
218. *Mycobacterium conspicuum*	NCBI	GCA_002102095.1_ASM210209v1
219. *Phenylobacterium* sp. Root700	NCBI	GCA_001429025.1_Root700
220. *Actinobacteria bacterium* IMCC26207	NCBI	GCA_001025035.1_ASM102503v1
221. *Verrucosispora* sp. SN26$_1$4.1	NCBI	GCA_004307965.1_ASM430796v1
222. *Sphingomonas* sp. Leaf412	NCBI	GCA_001425405.1_Leaf412
223. *Mycobacterium* sp. EPG1	NCBI	GCA_002946335.1_ASM294633v1
224. *Mycolicibacterium chubuense*	NCBI	GCA_000266905.1_ASM26690v1

Continued

Taxon Name	Database	Genome ID
225. *Chromatiales bacterium* (*Bugula neritina* AB1)	NCBI	GCA_002729495.1_ ASM272949v1
226. *Mycobacterium* sp. NAZ190054	NCBI	GCA_001545925.1_ASM154592v1
227. *Amycolatopsis circi*	NCBI	GCA_003385235.1_ASM338523v1
228. *Mycolicibacterium duvalii*	NCBI	GCA_002553585.1_ASM255358v1
229. *Gemmobacter intermedius*	NCBI	GCA_004054105.1_ASM405410v1
230. *Mycobacterium* sp. AT1	NCBI	GCA_002043095.1_ASM204309v1
231. *Sphingomonas* sp. JJ-A5 (*Tardibacter chloracetimidivorans*)	NCBI	GCA_001890385.1_ASM189038v1
232. *Mycobacterium* sp. Soil538	NCBI	GCA_001428285.1_Soil538
233. *Thalassococcus* sp. S3	NCBI	GCA_004216475.1_ASM421647v1
234. *Mycobacterium* sp. 852013-51886_SCH5428379	NCBI	GCA_001665575.1_ ASM166557v1
235. *Gemmobacter* sp. YIM 102744-1	NCBI	GCA_003863335.1_ASM386333v1
236. *Mycobacterium* sp. Root135	NCBI	GCA_001426545.1_Root135
237. *Candidatus Rhodobacter lobularis*	NCBI	GCA_001078595.1_ASM107859v1
238. *Sphingomonas* sp. TZW2008	NCBI	GCA_002117915.1_SpTZW2008_1.0
239. *Geodermatophilus* sp. Leaf369	NCBI	GCA_001424455.1_Leaf369
240. *Ktedonobacterales bacterium Uno16*	NCBI	GCA_003967575.1_ASM396757v1
241. *Ktedonobacterales bacterium Uno11*	NCBI	GCA_003967555.1_ASM396755v1
242. *Bradyrhizobium* sp. th.b2	NCBI	GCA_000426785.1_ASM42678v1
243. *Nostoc* sp. 3335mG	NCBI	GCA_003185865.1_ASM318586v1
244. *Mycobacterium* sp. CECT 8778	NCBI	GCA_002591975.1_ASM259197v1
245. *Ktedonobacterales bacterium* Uno3	NCBI	GCA_003967535.1_ASM396753v1
246. *Dehalococcoidia bacterium* SG8_51_3	NCBI	GCA_001303565.1_ASM130356v1
247. *Dehalococcoidia bacterium* SM23_28_2	NCBI	GCA_001303545.1_ASM130354v1
248. *Dehalococcoidia bacterium* CG2_30_46_9	NCBI	GCA_001873005.1_ ASM187300v1
249. `Dehalogenimonas alkenigignens`	NCBI	GCA_003095415.1_ASM309541v1
250. *Mycolicibacterium iranicum*	NCBI	GCA_002101705.1_ASM210170v1
251. *Salinispora arenicola* CNS-205 (*Micromonosporaceae*)	NCBI	GCA_000018265.1_ASM1826v1
252. *Gordonia* sp.	NCBI	GCA_003987705.1_ASM398770v1
253. *Albimonas pacifica*	NCBI	GCA_900113695.1_IMG-taxon_2667527406
254. *Alphaproteobacteria bacterium* PA2	NCBI	GCA_002256425.1_ASM225642v1
255. *Streptomycetaceae*	NCBI	GCA_000331185.1_ASM33118v1
256. *Saccharopolyspora flava*	NCBI	GCA_900116135.1_IMG-taxon_2599185150

Continued

Taxon Name	Database	Genome ID
257. *Bradyrhizobium japonicum*	NCBI	GCA_001887695.1_ASM188769v1
258. *Mycobacterium genavense*	NCBI	GCA_000526915.1_ASM52691v1
259. *Bradyrhizobium* sp. BTAi1	NCBI	GCA_000015165.1_ASM1516v1
260. *Dehalogenimonas formicexedens*	NCBI	GCA_001953175.1_ASM195317v1
261. *Mycolicibacterium rhodesiae*	NCBI	GCA_002086695.1_ASM208669v1
262. *Bradyrhizobium elkanii*	NCBI	GCA_001718185.1_ASM171818v1
263. *Bradyrhizobium* sp. STM 3843	NCBI	GCA_000239815.2_ASM23981v2
264. *Amycolatopsis australiensis*	NCBI	GCA_900119165.1_IMG-taxon_2642422517
265. *Bradyrhizobium valentinum*	NCBI	GCA_001440405.1_ASM144040v1
266. *Bradyrhizobium* sp. Tv2a-2	NCBI	GCA_000472425.1_ASM47242v1
267. *Saccharopolyspora shandongensis*	NCBI	GCA_900106995.1_IMG-taxon_2671180068
268. *Pseudonocardia* sp. CNS-004	NCBI	GCA_001942185.1_ASM194218v1
269. *Acidimicrobiales bacterium* mtb01	NCBI	GCA_004379135.1_ASM437913v1
270. *Minwuia thermotolerans*	NCBI	GCA_002924445.1_ASM292444v1
271. *Parahaliea mediterranea*	NCBI	GCA_003402235.1_ASM340223v1
272. *Rhizobium* sp. BK251	NCBI	GCA_004345245.1_ASM434524v1
273. *Rhizobium Actinomycetospora succinea*	NCBI	GCA_004363095.1_ASM436309v1
274. *Saccharopolyspora* sp. 7K502	NCBI	GCA_004348985.1_ASM434898v1
275. *Nocardia otitidiscaviarum*	NCBI	GCA_900454305.1_54984$_D$01
276. *Alphaproteobacteria bacterium* 62-8	NCBI	GCA_001898015.1_ASM189801v1
277. *Cumulibacter* sp. G-1	NCBI	GCA_004382795.1_ASM438279v1
278. *Gordonia westfalica*	NCBI	GCA_900105725.1_IMG-taxon_2636416072
279. *Pseudonocardia acaciae*	NCBI	GCA_000620785.1_ASM62078v1
280. *Actinopolyspora erythraea*	NCBI	GCA_002263515.1_ASM226351v1
281. *Sphingomonas* sp. SRS2	NCBI	GCA_000971055.1_SRS2_1.0
282. *Saccharothrix espanaensis*	NCBI	GCA_000328705.1_ASM32870v1
283. *Rhodopila globiformis*	NCBI	GCA_002937115.1_ASM293711v1
284. *Afipia* sp. P52-10	NCBI	GCA_000516555.1_P52-10
285. *Afipia broomeae*	NCBI	GCA_002797635.1_ASM279763v1
286. *Actinomycetospora chiangmaiensis*	NCBI	GCA_000379625.1_ASM37962v1
287. *Pseudonocardia endophytica*	NCBI	GCA_004339565.1_ASM433956v1
288. *Sphingomonas* sp. YZ-8	NCBI	GCA_003660165.1_ASM366016v1

Continued

Taxon Name	Database	Genome ID
289. *Pseudonocardia autotrophica*	NCBI	GCA_002119215.1_ASM211921v1
290. *Rhodococcus* sp. S2-17	NCBI	GCA_003130705.1_ASM313070v1
291. Nocardia niigatensis	NCBI	GCA_000308655.1_ASM30865v1
292. Dactylosporangium aurantiacum	NCBI	GCA_000716715.1_ASM71671v1
293. Allokutzneria albata	NCBI	GCA_900103775.1_IMG-taxon_2634166198
294. Streptomyces tsukubensis	NCBI	GCA_002007125.1_ASM200712v1
295. *Nocardia seriolae*	NCBI	GCA_001865855.1_ASM186585v1
296. Actinobacteria bacterium	NCBI	GCA_004376625.1_ASM437662v1
297. Enhydrobacter aerosaccus SK60	NCBI	GCA_000175915.1_ASM17591v1
298. Amycolatopsis sp. BJA-103	NCBI	GCA_002849735.1_ASM284973v1

Appendix B

Supplementary Information of Chapter 4

Let ($\chi^\star$, $\phi^\star$) be the fixed point of the dimensionless equation system (4.17) and (4.18), then we have

$$\xi\frac{\alpha+\beta\chi^\star}{\gamma+\chi^{\star 2}}=(\phi^\star+\nu)\frac{\chi^\star}{1+\chi^\star} \tag{B.1}$$

$$\xi\frac{\alpha+\beta\chi^\star}{\gamma+\chi^{\star 2}}=\phi^\star. \tag{B.2}$$

The RHS of (B.1) and (B.2) are equal. which gives

$$\phi^\star=\chi^\star\nu. \tag{B.3}$$

which can simplify the computation of the fixed point – we only need to find the $\chi^\star$ component of a fixed point and then can directly obtain $\phi^\star$. Substituting (B.3) to (4.18) gives

$$\xi\frac{\alpha+\beta\chi^\star}{\gamma+\chi^{\star 2}}=\nu\chi^\star, \tag{B.4}$$

which suggests a simple graphic identification of fixed points: the location where $y=\xi(\alpha+\beta\chi)/(\gamma+\chi^2)$ intersects $y=\nu\chi$.

Now we estimate the value of ν for which the system consisting of (4.17) and (4.18) changes its stability. We expect that the transition of the system's dynamical behaviors occurs as χ approaches $\mathcal{O}(1)$, where the Michaelis-Menten form $M(x)$ is at half of its maximum value. The amount of O_2 in modern Earth's atmosphere-ocean system has been suggested to be at a steady state (i.e, a fixed point) (Holland, 1978; Lyons et al., 2014), which gives the following relation according to (B.3):

$$\nu=\frac{\phi^\star}{\chi^\star}=\frac{y^\star}{x^\star}. \tag{B.5}$$

It has been estimated that the amount of organic carbon in modern Earth's crust is about $y^\star=1.5\times10^{21}$ mol and the amount of O_2 (and its oxidizing equivalent such as Fe_2O_3, SO_4^{2-}, etc.) in modern Earth's atmosphere, ocean, and sedimentary is about $x^\star=0.55\times10^{21}$ mol (Catling et al., 2001b). Thus, $\nu=y^\star/x^\star\approx2.73$. As discussed above, we expect that $\chi\sim\mathcal{O}(1)$, which implies that $x \sim k$ and therefore

$$v = \frac{k\nu}{t_y} \sim \frac{x\nu}{t_y}. \tag{B.6}$$

As discussed in Section 4, $t_y \approx 5 \times 10^8$ yr. Thus, $v = 3 \times 10^{12}$ mole/yr, which is consistent with estimation in Catling et al. (2001b). Also, with $\chi \sim \mathcal{O}(1)$, we expect

$$\nu \sim \mathcal{O}\left(\xi \frac{\alpha + \beta\chi^\star}{\gamma + \chi^{\star 2}}\right). \tag{B.7}$$

Using the results in (Hartnett et al., 1998) gives a rough value of the burial efficiency in the modern environment:

$$\frac{\alpha + \beta\chi^\star}{\gamma + \chi^{\star 2}} \sim \mathcal{O}(0.1). \tag{B.8}$$

Thus, we obtain

$$J_p \sim 10\nu \sim \mathcal{O}(10^{13}), \tag{B.9}$$

which is consistent with the estimated rate of organic carbon deposited on modern continental shelves is $J_p = 1.42 \times 10^{13}$ mole/yr (Galvez et al., 2020). These results show that the expectation that $\chi \sim \mathcal{O}(1)$ is reasonable.

Using the relation in (B.3), the solutions to the equation of (4.19) and (4.20) – that is, the fixed points – are

$$(\chi_1^\star, \phi_1^\star) = \left(\frac{\sqrt[3]{2}(c\nu - \beta\xi)}{S} - \frac{S}{3\sqrt[3]{2}\nu}, \quad \frac{\sqrt[3]{2}(\nu - \beta\xi)\nu}{S} - \frac{S}{3\sqrt[3]{2}} \right), \tag{B.10}$$

$$(\chi_2^\star, \phi_2^\star) = \left(\frac{\left(1 + i\sqrt{3}\right)(\beta\xi - \nu)}{2^{2/3}S} + \frac{\left(1 - i\sqrt{3}\right)S}{6\sqrt[3]{2}\nu}, \quad \frac{\left(1 + i\sqrt{3}\right)(\beta\xi - c\nu)\nu}{2^{2/3}S} + \frac{\left(1 - i\sqrt{3}\right)S}{6\sqrt[3]{2}} \right), \tag{B.11}$$

and

$$(\chi_3^\star, \phi_3^\star) = \left(\frac{\left(1 - i\sqrt{3}\right)(b\xi - \nu)}{2^{2/3}S} + \frac{\left(1 + i\sqrt{3}\right)S}{6\sqrt[3]{2}\nu}, \quad \frac{\left(1 - i\sqrt{3}\right)(b\xi - c\nu)\nu}{2^{2/3}S} + \frac{\left(1 + i\sqrt{3}\right)S}{6\sqrt[3]{2}} \right). \tag{B.12}$$

where

$$S = (R^{1/2} - 27a\xi\nu^2)^{1/3}, \tag{B.13}$$

and

$$R = 729\alpha^2\xi^2\nu^4 - 108\nu^3(\beta\xi - \nu)^3. \tag{B.14}$$

When $R < 0$, x_1 is a real number and x_2 and x_3 are two complex values. When $R = 0$, at least two solutions are equal. When $R > 0$, three solutions are real and different from one another.

Now we consider R as a function of ν.

$$R = 729\alpha^2\xi^2\nu^4 - 108\nu^3(\beta\xi - \nu)^3 = \nu^3[729\alpha^2\xi^2\nu - 108(\beta\xi - \nu)^3] \tag{B.15}$$

Rewrite the (4.17) and (4.18) as

$$\frac{d\chi}{d\tau} = F(\chi, \psi) \tag{B.16}$$

$$\frac{d\psi}{d\tau} = G(\chi, \psi). \tag{B.17}$$

We calculate the Jacobian at the fixed point ($\chi^{\star}$, $\phi^{\star}$):

$$J = \begin{pmatrix} \frac{\partial F}{\partial \chi} & \frac{\partial F}{\partial \psi} \\ \frac{\partial G}{\partial \chi} & \frac{\partial G}{\partial \psi} \end{pmatrix}_{(\chi=\chi^{\star}, \phi=\phi^{\star})} = \begin{pmatrix} \xi\frac{(1-\chi^{\star 2})\beta - 2\alpha\chi^{\star}}{(\gamma+\chi^{\star 2})^2} - \frac{\phi^{\star}+\nu}{(\gamma+\chi^{\star})^2} & -\frac{\chi^{\star}}{\gamma+\chi^{\star}} \\ \xi\frac{(1-\chi^{\star 2})\beta - 2\alpha\chi^{\star}}{(\gamma+\chi^{\star 2})^2} & -1 \end{pmatrix}. \tag{B.18}$$

Its determinant is

$$\Delta = -\xi\frac{(1-\chi^{\star 2})\beta - 2\alpha\chi^{\star}}{(\gamma+\chi^{\star 2})^2}\frac{1}{\gamma+\chi^{\star}} + \frac{\phi^{\star}+\nu}{(\gamma+\chi^{\star})^2} \tag{B.19}$$

and its trace is

$$\mathrm{Tr} = \xi\frac{(1-\chi^{\star 2})\beta - 2\alpha\chi^{\star}}{(\gamma+\chi^{\star 2})^2} - \frac{\nu}{\gamma+\chi^{\star}} - 1. \tag{B.20}$$

The stability of the steady state depends on the sign of the real part of the eigenvalues λ of the matrix $J = \lambda I$, where I is the identity matrix. The eigenvalues are given by

$$\lambda = \frac{\mathrm{Tr} \pm \sqrt{\mathrm{Tr}^2 - 4\Delta}}{2}. \tag{B.21}$$

The steady state of equation system (4.17) and (4.18) is stable when $\mathrm{Tr} < 0$ and unstable when $\mathrm{Tr} > 0$.

Bibliography

Alcott, L. J., Mills, B. J., and Poulton, S. W. (2019). Stepwise earth oxygenation is an inherent property of global biogeochemical cycling. *Science*, 366(6471):1333–1337.

Anbar, A. D. (2008). Elements and evolution. *Science*, 322(5907):1481–1483.

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(5846):1903–1906.

Arakawa, N., Aluwihare, L. I., Simpson, A. J., Soong, R., Stephens, B. M., and Lane-Coplen, D. (2017). Carotenoids are the likely precursor of a significant fraction of marine dissolved organic matter. *Science advances*, 3(9):e1602976.

Arnarson, T. S. and Keil, R. G. (2007a). Changes in organic matter–mineral interactions for marine sediments with varying oxygen exposure times. *Geochimica et Cosmochimica Acta*, 71(14):3545–3556.

Arnarson, T. S. and Keil, R. G. (2007b). Changes in organic matter–mineral interactions for marine sediments with varying oxygen exposure times. *Geochimica et Cosmochimica Acta*, 71(14):3545–3556.

Arndt, S., Jørgensen, B. B., LaRowe, D. E., Middelburg, J., Pancost, R., and Regnier, P. (2013). Quantifying the degradation of organic matter in marine sediments: a review and synthesis. *Earth-science reviews*, 123:53–86.

Ataman, E., Andersson, M. P., Ceccato, M., Bovet, N., and Stipp, S. L. S. (2016). Functional group adsorption on calcite: I. oxygen containing and nonpolar organic molecules. *The Journal of Physical Chemistry C*, 120(30):16586–16596.

Aulbach, S. and Stagno, V. (2016). Evidence for a reducing archean ambient mantle and its effects on the carbon cycle. *Geology*, 44(9):751–754.

Bachan, A. and Kump, L. R. (2015). The rise of oxygen and siderite oxidation during the lomagundi event. *Proceedings of the National Academy of Sciences*, 112(21):6562–6567.

Baldock, J. and Skjemstad, J. (2000). Role of the soil matrix and minerals in protecting natural organic materials against biological attack. *Organic geochemistry*, 31(7-8):697–710.

Bansal, M. S., Kellis, M., Kordi, M., and Kundu, S. (2018). RANGER-DTL 2.0: rigorous reconstruction of gene-family evolution by duplication, transfer and loss. *Bioinformatics*, 34(18):3214–3216.

Barber, A., Brandes, J., Leri, A., Lalonde, K., Balind, K., Wirick, S., Wang, J., and Gélinas, Y. (2017). Preservation of organic matter in marine sediments by inner-sphere interactions with reactive iron. *Scientific reports*, 7(1):1–10.

Battistuzzi, F. U. and Hedges, S. B. (2009). A major clade of prokaryotes with ancient adaptations to life on land. *Molecular biology and evolution*, 26(2):335–343.

Benner, R., Pakulski, J. D., McCarthy, M., Hedges, J. I., and Hatcher, P. G. (1992). Bulk chemical characteristics of dissolved organic matter in the ocean. *Science*, 255(5051):1561–1564.

Berner, R. A. (1964). An idealized model of dissolved sulfate distribution in recent sediments. *Geochimica et Cosmochimica Acta*, 28(9):1497–1503.

Berner, R. A. (1990). Atmospheric carbon dioxide levels over phanerozoic time. *Science*, 249(4975):1382–1386.

Berner, R. A. (1999). Atmospheric oxygen over phanerozoic time. *Proceedings of the National Academy of Sciences*, 96(20):10955–10957.

Berner, R. A. (2004). *The Phanerozoic carbon cycle: CO2 and O2*. Oxford University Press.

Blair, N. E. and Aller, R. C. (2012). The fate of terrestrial organic carbon in the marine environment. *Annual review of marine science*, 4:401–423.

Blattmann, T. M., Liu, Z., Zhang, Y., Zhao, Y., Haghipour, N., Montluçon, D. B., Plötze, M., and Eglinton, T. I. (2019). Mineralogical control on the fate of continentally derived organic matter in the ocean. *Science*, 366(6466):742–745.

Boudreau, B. P. and Ruddick, B. R. (1991). On a reactive continuum representation of organic matter diagenesis. *American Journal of Science*, 291(5):507–538.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Boye, K., Noël, V., Tfaily, M. M., Bone, S. E., Williams, K. H., Bargar, J. R., and Fendorf, S. (2017). Thermodynamically controlled preservation of organic carbon in floodplains. *Nature Geoscience*, 10(6):415–419.

Brenner, A. R., Fu, R. R., Evans, D. A., Smirnov, A. V., Trubko, R., and Rose, I. R. (2020). Paleomagnetic evidence for modern-like plate motion velocities at 3.2 ga. *Science Advances*, 6(17):eaaz8670.

Buick, R. (2008). When did oxygenic photosynthesis evolve? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1504):2731–2743.

Burdige, D. J. (2007). Preservation of organic matter in marine sediments: controls, mechanisms, and an imbalance in sediment organic carbon budgets? *Chemical reviews*, 107(2):467–485.

Campbell, I. H. and Allen, C. M. (2008). Formation of supercontinents linked to increases in atmospheric oxygen. *Nature Geoscience*, 1(8):554–558.

Canfield, D. E. (1994). Factors influencing organic carbon preservation in marine sediments. *Chemical geology*, 114(3-4):315–329.

Canfield, D. E. (2014). *Oxygen: a four billion year history*. Princeton University Press.

Canfield, D. E., Thamdrup, B., and Hansen, J. W. (1993). The anaerobic degradation of organic matter in danish coastal sediments: iron reduction, manganese reduction, and sulfate reduction. *Geochimica et Cosmochimica Acta*, 57(16):3867–3883.

Cao, X., Aiken, G. R., Butler, K. D., Huntington, T. G., Balch, W. M., Mao, J., and Schmidt-Rohr, K. (2018). Evidence for major input of riverine organic matter into the ocean. *Organic Geochemistry*, 116:62–76.

Catling, D. C. and Claire, M. W. (2005). How earth's atmosphere evolved to an oxic state: a status report. *Earth and Planetary Science Letters*, 237(1-2):1–20.

Catling, D. C. and Kasting, J. F. (2017). *Atmospheric evolution on inhabited and lifeless worlds*. Cambridge University Press.

Catling, D. C., Zahnle, K. J., and McKay, C. P. (2001a). Biogenic methane, hydrogen escape, and the irreversible oxidation of early earth. *Science*, 293(5531):839–843.

Catling, D. C., Zahnle, K. J., and McKay, C. P. (2001b). Biogenic methane, hydrogen escape, and the irreversible oxidation of early earth. *Science*, 293(5531):839–843.

Chassé, A. W., Ohno, T., Higgins, S. R., Amirbahman, A., Yildirim, N., and Parr, T. B. (2015). Chemical force spectroscopy evidence supporting the layer-by-layer model of organic matter binding to iron (oxy) hydroxide mineral surfaces. *Environmental science & technology*, 49(16):9733–9741.

Cox, G. M., Halverson, G. P., Stevenson, R. K., Vokaty, M., Poirier, A., Kunzmann, M., Li, Z.-X., Denyszyn, S. W., Strauss, J. V., and Macdonald, F. A. (2016). Continental flood basalt weathering as a trigger for neoproterozoic snowball earth. *Earth and Planetary Science Letters*, 446:89–99.

Dasgupta, R. and Hirschmann, M. M. (2010). The deep carbon cycle and melting in earth's interior. *Earth and Planetary Science Letters*, 298(1-2):1–13.

Davis, J. A. (1982). Adsorption of natural dissolved organic matter at the oxide/water interface. *Geochimica et Cosmochimica Acta*, 46(11):2381–2393.

DeLapp, R. C., LeBoeuf, E. J., and Bell, K. D. (2004). Thermodynamic properties of several soil-and sediment-derived natural organic materials. *Chemosphere*, 54(4):527–539.

Deshmukh, A. P., Pacheco, C., Hay, M. B., and Myneni, S. C. (2007). Structural environments of carboxyl groups in natural organic molecules from terrestrial systems. part 2: 2d nmr spectroscopy. *Geochimica et cosmochimica acta*, 71(14):3533–3544.

Dhooge, A., Govaerts, W., and Kuznetsov, Y. A. (2003). Matcont: a matlab package for numerical bifurcation analysis of odes. *ACM Transactions on Mathematical Software (TOMS)*, 29(2):141–164.

DiDonato, N., Chen, H., Waggoner, D., and Hatcher, P. G. (2016). Potential origin and formation for molecular components of humic acids in soils. *Geochimica et Cosmochimica acta*, 178:210–222.

Drouin, S., Boussafir, M., Robert, J.-L., Albéric, P., and Durand, A. (2010). Carboxylic acid sorption on synthetic clays in sea water: in vitro experiments and implications for organo-clay behaviour under marine conditions. *Organic Geochemistry*, 41(2):192–199.

Duncan, M. S. and Dasgupta, R. (2017). Rise of earth's atmospheric oxygen controlled by efficient subduction of organic carbon. *Nature Geoscience*, 10(5):387–392.

Eguchi, J., Seales, J., and Dasgupta, R. (2020). Great oxidation and lomagundi events linked by deep cycling and enhanced degassing of carbon. *Nature geoscience*, 13(1):71–76.

Emerson, S. and Hedges, J. (1988). Processes controlling the organic carbon content of open ocean sediments. *Paleoceanography*, 3(5):621–634.

Fairchild, I. J., Spencer, A. M., Ali, D. O., Anderson, R. P., Anderton, R., Boomer, I., Dove, D., Evans, J. D., Hambrey, M. J., Howe, J., et al. (2018). Tonian-cryogenian boundary sections of argyll, scotland. *Precambrian Research*, 319:37–64.

Farquhar, J., Zerkle, A. L., and Bekker, A. (2011). Geological constraints on the origin of oxygenic photosynthesis. *Photosynthesis research*, 107(1):11–36.

Fava, A. and Eyring, H. (1956). Equilibrium and kinetics of detergent adsorption–a generalized equilibration theory. *The Journal of Physical Chemistry*, 60(7):890–898.

Fike, D., Grotzinger, J., Pratt, L., and Summons, R. (2006). Oxidation of the ediacaran ocean. *nature*, 444(7120):744–747.

Fischer, R. A., Cottrell, E., Hauri, E., Lee, K. K., and Le Voyer, M. (2020). The carbon content of earth and its core. *Proceedings of the National Academy of Sciences*, 117(16):8743–8749.

Forney, D. C. and Rothman, D. H. (2012). Common structure in the heterogeneity of plant-matter decay. *Journal of the Royal Society Interface*, 9(74):2255–2267.

Fournier, G. P., Huang, J., and Gogarten, J. P. (2009). Horizontal gene transfer from extinct and extant lineages: biological innovation and the coral of life. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1527):2229–2239.

Frantz, P., Granick, S., Iyengar, D., and McCarthy, T. J. (1990). End group effect in polymer adsorption: Competitive adsorption of carboxylic acid-terminated and unfunctionalized polystyrene. *The Journal of chemical physics*, 92(11):6970–6971.

Fuchs, G., Boll, M., and Heider, J. (2011). Microbial degradation of aromatic compounds—from one strategy to four. *Nature Reviews Microbiology*, 9(11):803–816.

Gaillard, F., Scaillet, B., and Arndt, N. T. (2011). Atmospheric oxygenation caused by a change in volcanic degassing pressure. *Nature*, 478(7368):229–232.

Galvez, M. E., Fischer, W. W., Jaccard, S. L., and Eglinton, T. I. (2020). Materials and pathways of the organic carbon cycle through time. *Nature geoscience*, pages 1–12.

Goldblatt, C., Lenton, T. M., and Watson, A. J. (2006). Bistability of atmospheric oxygen and the great oxidation. *Nature*, 443(7112):683–686.

Gu, B., Schmitt, J., Chen, Z., Liang, L., and McCarthy, J. F. (1994). Adsorption and desorption of natural organic matter on iron oxide: mechanisms and models. *Environmental Science & Technology*, 28(1):38–46.

Guckenheimer, J. and Holmes, P. (2013). *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Springer Science & Business Media.

Halverson, G. P. (2006). A neoproterozoic chronology. In *Neoproterozoic geobiology and paleobiology*, pages 231–271. Springer.

Halverson, G. P., Hoffman, P. F., Schrag, D. P., Maloof, A. C., and Rice, A. H. N. (2005). Toward a neoproterozoic composite carbon-isotope record. *Geological Society of America Bulletin*, 117(9-10):1181–1207.

Harayama, S., Kok, M., and Neidle, E. (1992). Functional and evolutionary relationships among diverse oxygenases. *Annual review of microbiology*, 46(1):565–601.

Hartnett, H. E., Keil, R. G., Hedges, J. I., and Devol, A. H. (1998). Influence of oxygen exposure time on organic carbon preservation in continental margin sediments. *Nature*, 391(6667):572–575.

Hay, M. B. and Myneni, S. C. (2007). Structural environments of carboxyl groups in natural organic molecules from terrestrial systems. part 1: Infrared spectroscopy. *Geochimica et Cosmochimica Acta*, 71(14):3518–3532.

Hayes, J. M. and Waldbauer, J. R. (2006). The carbon cycle and associated redox processes through time. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1470):931–950.

Hazen, R. M. and Ferry, J. M. (2010). Mineral evolution: Mineralogy in the fourth dimension. *Elements*, 6(1):9–12.

Hazen, R. M., Papineau, D., Bleeker, W., Downs, R. T., Ferry, J. M., McCoy, T. J., Sverjensky, D. A., and Yang, H. (2008). Mineral evolution. *American Mineralogist*, 93(11-12):1693–1720.

Hazen, R. M., Sverjensky, D. A., Azzolini, D., Bish, D. L., Elmore, S. C., Hinnov, L., and Milliken, R. E. (2013). Clay mineral evolution. *American Mineralogist*, 98(11-12):2007–2029.

Heckman, D. S., Geiser, D. M., Eidell, B. R., Stauffer, R. L., Kardos, N. L., and Hedges, S. B. (2001). Molecular evidence for the early colonization of land by fungi and plants. *science*, 293(5532):1129–1133.

Hedges, J. (2002). Sedimentary organic matter preservation and atmospheric O_2 regulation. In *Chemistry of marine water and sediments*, pages 105–123. Springer.

Hedges, J. I., Hu, F. S., Devol, A. H., Hartnett, H. E., Tsamakis, E., and Keil, R. G. (1999). Sedimentary organic matter preservation; a test for selective degradation under oxic conditions. *American Journal of Science*, 299(7-9):529–555.

Hedges, J. I. and Keil, R. G. (1995). Sedimentary organic matter preservation: an assessment and speculative synthesis. *Marine chemistry*, 49(2-3):81–115.

Hemingway, J. D., Rothman, D. H., Grant, K. E., Rosengard, S. Z., Eglinton, T. I., Derry, L. A., and Galy, V. V. (2019). Mineral protection regulates long-term global preservation of natural organic carbon. *Nature*, 570(7760):228–231.

Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., and Hedges, J. I. (2006). Characterization of a major refractory component of marine dissolved organic matter. *Geochimica et Cosmochimica Acta*, 70(12):2990–3010.

Hertkorn, N., Ruecker, C., Meringer, M., Gugisch, R., Frommberger, M., Perdue, E., Witt, M., and Schmitt-Kopplin, P. (2007). High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. *Analytical and bioanalytical chemistry*, 389(5):1311–1327.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). Ufboot2: improving the ultrafast bootstrap approximation. *Molecular biology and evolution*, 35(2):518–522.

Hoehler, T. (2004). Biological energy requirements as quantitative boundary conditions for life in the subsurface. *Geobiology*, 2(4):205–215.

Hoffman, P. F., Abbot, D. S., Ashkenazy, Y., Benn, D. I., Brocks, J. J., Cohen, P. A., Cox, G. M., Creveling, J. R., Donnadieu, Y., Erwin, D. H., et al. (2017). Snowball Earth climate dynamics and cryogenian geology-geobiology. *Science Advances*, 3(11):e1600983.

Holland, H. D. (1978). *The chemistry of the atmosphere and oceans*. John Wiley Sons.

Holland, H. D. (2002). Volcanic gases, black smokers, and the great oxidation event. *Geochimica et Cosmochimica Acta*, 66(21):3811–3826.

Holland, H. D. (2006). The oxygenation of the atmosphere and oceans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1470):903–915.

Holland, H. D. (2009). Why the atmosphere became oxygenated: a proposal. *Geochimica et Cosmochimica Acta*, 73(18):5241–5255.

Huang, W., Yu, Z., Fu, J., et al. (2003). Effects of organic matter heterogeneity on sorption and desorption of organic contaminants by soils and sediments. *Applied Geochemistry*, 18(7):955–972.

Izhikevich, E. M. (2007). *Dynamical systems in neuroscience*. MIT press.

Jenkyns, H. C. (2010). Geochemistry of oceanic anoxic events. *Geochemistry, Geophysics, Geosystems*, 11(3).

Jin, Q. and Bethke, C. M. (2005). Predicting the rate of microbial respiration in geochemical environments. *Geochimica et Cosmochimica Acta*, 69(5):1133–1143.

Jinich, A., Sanchez-Lengeling, B., Ren, H., Goldford, J. E., Noor, E., Sanders, J. N., Segrè, D., and Aspuru-Guzik, A. (2020). A thermodynamic atlas of carbon redox chemical space. *Proceedings of the National Academy of Sciences*, 117(52):32910–32918.

Jørgensen, B. (1978). Comparison of methods for the quantification of bacterial sulfate reduction in coastal marine-sediments. 2. calculation from mathematical-models. *Geomicrobiology Journal*, 1(1):29–47.

Kacar, B., Guy, L., Smith, E., and Baross, J. (2017). Resurrecting ancestral genes in bacteria to interpret ancient biosignatures. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2109):20160352.

Kadoya, S., Catling, D., Nicklas, R., Puchtel, I., and Anbar, A. (2020a). Mantle cooling causes more reducing volcanic gases and gradual reduction of the atmosphere. *Geochem. Perspect. Lett.*, 13:25–29.

Kadoya, S., Catling, D. C., Nicklas, R. W., Puchtel, I. S., and Anbar, A. D. (2020b). Mantle data imply a decline of oxidizable volcanic gases could have triggered the great oxidation. *Nature Communications*, 11(1):1–9.

Kaiser, K. and Guggenberger, G. (2000). The role of dom sorption to mineral surfaces in the preservation of organic matter in soils. *Organic geochemistry*, 31(7-8):711–725.

Kasting, J. F. (1987). Theoretical constraints on oxygen and carbon dioxide concentrations in the precambrian atmosphere. *Precambrian research*, 34(3-4):205–229.

Kasting, J. F. (2001). The rise of atmospheric oxygen. *Science*, 293(5531):819–820.

Kasting, J. F. (2013). What caused the rise of atmospheric o2? *Chemical Geology*, 362:13–25.

Katoh, K. and Standley, D. M. (2013). Mafft multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4):772–780.

Kaufman, A. J., Johnston, D. T., Farquhar, J., Masterson, A. L., Lyons, T. W., Bates, S., Anbar, A. D., Arnold, G. L., Garvin, J., and Buick, R. (2007). Late archean biospheric oxygenation and atmospheric evolution. *Science*, 317(5846):1900–1903.

Kearey, P., Klepeis, K. A., and Vine, F. J. (2009). *Global tectonics*. John Wiley & Sons.

Keil, R. G. and Mayer, L. M. (2014). Mineral matrices and organic matter. In Holland, H. D. and Turekian, K. K., editors, *Treatise on Geochemistry*, volume 12, pages 337–359. Elsevier, Oxford, 2 edition.

Keil, R. G., Montluçon, D. B., Prahl, F. G., and Hedges, J. I. (1994). Sorptive preservation of labile organic matter in marine sediments. *Nature*, 370(6490):549–552.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015). The phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, 10(6):845.

Kennedy, M., Droser, M., Mayer, L. M., Pevear, D., and Mrofka, D. (2006). Late precambrian oxygenation; inception of the clay mineral factory. *Science*, 311(5766):1446–1449.

Kennedy, M. J., Pevear, D. R., and Hill, R. J. (2002). Mineral surface control of organic carbon in black shale. *Science*, 295(5555):657–660.

Kile, D. E., Wershaw, R. L., and Chiou, C. T. (1999). Correlation of soil and sediment organic matter polarity to aqueous sorption of nonionic compounds. *Environmental science & technology*, 33(12):2053–2056.

Kim, K. M., Qin, T., Jiang, Y.-Y., Chen, L.-L., Xiong, M., Caetano-Anollés, D., Zhang, H.-Y., and Caetano-Anollés, G. (2012). Protein domain structure uncovers the origin of aerobic metabolism and the rise of planetary oxygen. *Structure*, 20(1):67–76.

Klaas, C. and Archer, D. E. (2002). Association of sinking organic matter with various types of mineral ballast in the deep sea: Implications for the rain ratio. *Global Biogeochemical Cycles*, 16(4):63–1.

Knoll, A. H. (2003). Biomineralization and evolutionary history. *Reviews in mineralogy and geochemistry*, 54(1):329–356.

Knoll, A. H. (2014). Paleobiological perspectives on early eukaryotic evolution. *Cold Spring Harbor Perspectives in Biology*, 6(1):a016121.

Knoll, A. H., Kaufman, A. J., and Semikhatov, M. A. (1995). The carbon-isotopic composition of proterozoic carbonates: Riphean successions from northwestern siberia (anabar massif, turukhansk uplift). *American Journal of Science*, 295(7):823–850.

Krissansen-Totton, J., Buick, R., and Catling, D. C. (2015). A statistical analysis of the carbon isotope record from the archean to phanerozoic and implications for the rise of oxygen. *American Journal of Science*, 315(4):275–316.

Kristensen, E., Ahmed, S. I., and Devol, A. H. (1995). Aerobic and anaerobic decomposition of organic matter in marine sediment: which is fastest? *Limnology and oceanography*, 40(8):1430–1437.

Kump, L. R. and Barley, M. E. (2007). Increased subaerial volcanism and the rise of atmospheric oxygen 2.5 billion years ago. *Nature*, 448(7157):1033–1036.

Kump, L. R., Kasting, J. F., and Barley, M. E. (2001). Rise of atmospheric oxygen and the "upside-down" archean mantle. *Geochemistry, Geophysics, Geosystems*, 2(1).

Kuznetsov, A., Semikhatov, M., Maslov, A., Gorokhov, I., Prasolov, E., Krupenin, M., and Kislova, I. (2006). New data on sr-and c-isotopic chemostratigraphy of the upper riphean type section (southern urals). *Stratigraphy and Geological Correlation*, 14(6):602–628.

Laakso, T. A. and Schrag, D. (2017). A theory of atmospheric oxygen. *Geobiology*, 15(3):366–384.

Lalonde, K., Mucci, A., Ouellet, A., and Gélinas, Y. (2012). Preservation of organic matter in sediments promoted by iron. *Nature*, 483(7388):198–200.

Landry, Z., Swan, B. K., Herndl, G. J., Stepanauskas, R., and Giovannoni, S. J. (2017). SAR202 genomes from the dark ocean predict pathways for the oxidation of recalcitrant dissolved organic matter. *MBio*, 8(2):e00413–17.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., et al. (2007). Clustal w and clustal x version 2.0. *bioinformatics*, 23(21):2947–2948.

LaRowe, D. E. and Van Cappellen, P. (2011). Degradation of natural organic matter: a thermodynamic analysis. *Geochimica et Cosmochimica Acta*, 75(8):2030–2042.

Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*, 25(17):2286–2288.

Le, S. Q. and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Molecular biology and evolution*, 25(7):1307–1320.

Lechtenfeld, O. J., Hertkorn, N., Shen, Y., Witt, M., and Benner, R. (2015). Marine sequestration of carbon in bacterial metabolites. *Nature communications*, 6:6711.

Leenheer, J. (1994). Chemistry of dissolved organic matter in rivers, lakes, and reservoirs. ACS Publications.

Lemey, P., Salemi, M., and Vandamme, A.-M. (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press.

Lenton, T. M. (1998). Gaia and natural selection. *Nature*, 394(6692):439–447.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J. (2008). Tipping elements in the Earth's climate system. *Proceedings of the national Academy of Sciences*, 105(6):1786–1793.

Liu, S., Parsons, R., Opalk, K., Baetge, N., Giovannoni, S., Bolaños, L. M., Kujawinski, E. B., Longnecker, K., Lu, Y., Halewood, E., et al. (2020). Different carboxyl-rich alicyclic molecules proxy compounds select distinct bacterioplankton for oxidation of dissolved organic matter in the mesopelagic sargasso sea. *Limnology and Oceanography*, 65(7):1532–1553.

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(5):e1600134.

Lyons, T. W., Reinhard, C. T., and Planavsky, N. J. (2014). The rise of oxygen in earth's early ocean and atmosphere. *Nature*, 506(7488):307–315.

Macdonald, F., Halverson, G., Strauss, J., Smith, E., Cox, G., Sperling, E., and Roots, C. (2012). Early neoproterozoic basin formation in yukon, canada:: Implications for the make-up and break-up of rodinia. *Geoscience Canada*, 39(2):77–100.

Macouin, M., Roques, D., Rousse, S., Ganne, J., Denele, Y., and Trindade, R. Y. (2015). Is the neoproterozoic oxygen burst a supercontinent legacy? *Frontiers in Earth Science*, 3:44.

Madsen, L., Grahl-Madsen, L., Grøn, C., Lind, I., and Engell, J. (1996). Adsorption of polar aromatic hydrocarbons on synthetic calcite. *Organic Geochemistry*, 24(12):1151–1155.

Magnabosco, C., Moore, K. R., Wolfe, J. M., and Fournier, G. P. (2018). Dating phototrophic microbial lineages with reticulate gene histories. *Geobiology*, 16(2):179–189.

Maloof, A. C., Porter, S. M., Moore, J. L., Dudás, F. Ö., Bowring, S. A., Higgins, J. A., Fike, D. A., and Eddy, M. P. (2010). The earliest cambrian record of animals and ocean geochemical change. *Bulletin*, 122(11-12):1731–1774.

Martin, A. P., Condon, D. J., Prave, A. R., and Lepland, A. (2013). A review of temporal constraints for the palaeoproterozoic large, positive carbonate carbon isotope excursion (the Lomagundi–Jatuli Event). *Earth-Science Reviews*, 127:242–261.

Mayer, L. M. (1994). Surface area control of organic carbon accumulation in continental shelf sediments. *Geochimica et Cosmochimica Acta*, 58(4):1271–1284.

Mayer, L. M. (2004). The inertness of being organic. *Marine Chemistry*, 92(1-4):135–140.

McKirdy, D. M., Burgess, J. M., Lemon, N. M., Yu, X., Cooper, A. M., Gostin, V. A., Jenkins, R. J., and Both, R. A. (2001). A chemostratigraphic overview of the late cryogenian interglacial sequence in the adelaide fold-thrust belt, south australia. *Precambrian Research*, 106(1-2):149–186.

Mehrshad, M., Rodriguez-Valera, F., Amoozegar, M. A., López-García, P., and Ghai, R. (2018). The enigmatic sar202 cluster up close: shedding light on a globally distributed dark ocean lineage involved in sulfur cycling. *The ISME journal*, 12(3):655–668.

Middelburg, J. J. (1989). A simple rate model for organic matter decomposition in marine sediments. *Geochimica et Cosmochimica Acta*, 53(7):1577–1581.

Middelburg, J. J. and Meysman, F. J. (2007). Burial at sea. *Science*, 316(5829):1294–1295.

Morris, R., Rappe, M., Urbach, E., Connon, S., and Giovannoni, S. (2004). Prevalence of the chloroflexi-related sar202 bacterioplankton cluster throughout the mesopelagic zone and deep ocean. *Applied and Environmental Microbiology*, 70(5):2836–2842.

Nebbioso, A. and Piccolo, A. (2013). Molecular characterization of dissolved organic matter (dom): a critical review. *Analytical and Bioanalytical Chemistry*, 405(1):109–124.

Nelson, D. L., Lehninger, A. L., and Cox, M. M. (2008). *Lehninger principles of biochemistry*. Macmillan, New York.

Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution*, 32(1):268–274.

Nicklas, R. W., Puchtel, I. S., Ash, R. D., Piccoli, P. M., Hanski, E., Nisbet, E. G., Waterton, P., Pearson, D. G., and Anbar, A. D. (2019). Secular mantle oxidation across the archean-proterozoic boundary: Evidence from v partitioning in komatiites and picrites. *Geochimica et Cosmochimica Acta*, 250:49–75.

Oades, J. (1988). The retention of organic matter in soils. *Biogeochemistry*, 5(1):35–70.

Olson, S. L., Jansen, M., and Abbot, D. S. (2020). Oceanographic considerations for exoplanet life detection. *The Astrophysical Journal*, 895(1):19.

Peterson, C. and Kwei, T. (1961). The kinetics of polymer adsorption onto solid surfaces. *The Journal of Physical Chemistry*, 65(8):1330–1333.

Plonka, A. (2013). *Dispersive kinetics*. Springer Science & Business Media.

Popper, K. (1959). *The logic of scientific discovery*. Routledge.

Prigogine, I. (1980). *From being to becoming: time and complexity in the physical sciences*. Freeman.

Ransom, B., Kim, D., Kastner, M., and Wainwright, S. (1998). Organic matter preservation on continental slopes: importance of mineralogy and surface area. *Geochimica et Cosmochimica Acta*, 62(8):1329–1345.

Rashid, M., Buckley, D., and Robertson, K. (1972). Interactions of a marine humic acid with clay minerals and a natural sediment. *Geoderma*, 8(1):11–27.

Rashid, M. A. (2012). *Geochemistry of marine humic compounds*. Springer Science & Business Media, New York.

Repeta, D. J. (2015). Chemical characterization and cycling of dissolved organic matter. In Hansell, D. A. and Carlson, C. A., editors, *Biogeochemistry of marine dissolved organic matter*, pages 21–63. Academic Press, Boston, 2 edition.

Rooney, A. D., Macdonald, F. A., Strauss, J. V., Dudás, F. Ö., Hallmann, C., and Selby, D. (2014). Re-os geochronology and coupled os-sr isotope constraints on the sturtian snowball earth. *Proceedings of the National Academy of Sciences*, 111(1):51–56.

Ross, J. and Vlad, M. O. (1999). Nonlinear kinetics and new approaches to complex reaction mechanisms. *Annual review of physical chemistry*, 50(1):51–78.

Rothman, D. (2015). Earth's carbon cycle: a mathematical perspective. *Bulletin of the American Mathematical Society*, 52(1):47–64.

Rothman, D. H. and Forney, D. C. (2007). Physical model for the decay and preservation of marine organic carbon. *Science*, 316(5829):1325–1328.

Sánchez-Baracaldo, P., Raven, J. A., Pisani, D., and Knoll, A. H. (2017). Early photosynthetic eukaryotes inhabited low-salinity habitats. *Proceedings of the National Academy of Sciences*, 114(37):E7737–E7745.

Santha, N., Cubillas, P., Saw, A., Brooksbank, H., and Greenwell, H. C. (2017). Chemical force microscopy study on the interactions of COOH functional groups with kaolinite surfaces: Implications for enhanced oil recovery. *Minerals*, 7(12):250.

Schlautman, M. A. and Morgan, J. J. (1994). Adsorption of aquatic humic substances on colloidal-size aluminum oxide particles: influence of solution chemistry. *Geochimica et Cosmochimica Acta*, 58(20):4293–4303.

Schmidt, M. W., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A., et al. (2011). Persistence of soil organic matter as an ecosystem property. *Nature*, 478(7367):49.

Schopf, J. W. (1993). Microfossils of the early archean apex chert: new evidence of the antiquity of life. *Science*, 260(5108):640–646.

Schrenk, M. O., Huber, J. A., and Edwards, K. J. (2010). Microbial provinces in the subseafloor. *Annual review of marine science*, 2:279–304.

Sessions, A. L., Doughty, D. M., Welander, P. V., Summons, R. E., and Newman, D. K. (2009). The continuing puzzle of the great oxidation event. *Current Biology*, 19(14):R567–R574.

Simberloff, D. (2014). The "balance of nature"—evolution of a panchreston. *PLoS Biology*, 12(10):e1001963.

Slotznick, S. P., Swanson-Hysell, N. L., and Sperling, E. A. (2018). Oxygenated mesoproterozoic lake revealed through magnetic mineralogy. *Proceedings of the National Academy of Sciences*, 115(51):12938–12943.

Smith, R. W., Bianchi, T. S., Allison, M., Savage, C., and Galy, V. (2015). High rates of organic carbon burial in fjord sediments globally. *Nature Geoscience*, 8(6):450–453.

Sobek, S., Durisch-Kaiser, E., Zurbrügg, R., Wongfun, N., Wessels, M., Pasche, N., and Wehrli, B. (2009). Organic carbon burial efficiency in lake sediments controlled by oxygen exposure time and sediment source. *Limnology and Oceanography*, 54(6):2243–2254.

Soo, R. M., Hemp, J., Parks, D. H., Fischer, W. W., and Hugenholtz, P. (2017). On the origins of oxygenic photosynthesis and aerobic respiration in cyanobacteria. *Science*, 355(6332):1436–1440.

Soubrier, J., Steel, M., Lee, M. S., Der Sarkissian, C., Guindon, S., Ho, S. Y., and Cooper, A. (2012). The influence of rate heterogeneity among sites on the time dependence of molecular rates. *Molecular Biology and Evolution*, 29(11):3345–3358.

Soucy, S. M., Huang, J., and Gogarten, J. P. (2015). Horizontal gene transfer: building the web of life. *Nature Reviews Genetics*, 16(8):472–482.

Strogatz, S. H. (2018). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. CRC press, New York.

Stüeken, E. E., Buick, R., and Anbar, A. D. (2015). Selenium isotopes support free o2 in the latest archean. *Geology*, 43(3):259–262.

Sverjensky, D. A. and Lee, N. (2010). The great oxidation event and mineral diversification. *Elements*, 6(1):31–36.

Taverne, Y. J., Caron, A., Diamond, C., Fournier, G., and Lyons, T. W. (2020). Oxidative stress and the early coevolution of life and biospheric oxygen. In *Oxidative Stress*, pages 67–85. Elsevier.

Tipping, E. (1981). The adsorption of aquatic humic substances by iron oxides. *Geochimica et Cosmochimica Acta*, 45(2):191–199.

Tolmie, C., Smit, M. S., and Opperman, D. J. (2019). Native roles of baeyer–villiger monooxygenases in the microbial metabolism of natural compounds. *Natural product reports*, 36(2):326–353.

Tria, F. D. K., Landan, G., and Dagan, T. (2017). Phylogenetic rooting using minimal ancestor deviation. *Nature ecology & evolution*, 1(1):1–7.

Vaidya, G., Lohman, D. J., and Meier, R. (2011). Sequencematrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, 27(2):171–180.

Valley, J. W., Cavosie, A. J., Ushikubo, T., Reinhard, D. A., Lawrence, D. F., Larson, D. J., Clifton, P. H., Kelly, T. F., Wilde, S. A., Moser, D. E., et al. (2014). Hadean age for a post-magma-ocean zircon confirmed by atom-probe tomography. *Nature Geoscience*, 7(3):219–223.

Vandenbroucke, M. and Largeau, C. (2007). Kerogen origin, evolution and structure. *Organic Geochemistry*, 38(5):719–833.

Vlad, M. O., Huber, D. L., and Ross, J. (1997). Rate statistics and thermodynamic analogies for relaxation processes in systems with static disorder: Application to stretched exponential. *The Journal of chemical physics*, 106(10):4157–4167.

Wagai, R. and Mayer, L. M. (2007). Sorptive stabilization of organic matter in soils by hydrous iron oxides. *Geochimica et Cosmochimica Acta*, 71(1):25–35.

Waggoner, D. C., Wozniak, A. S., Cory, R. M., and Hatcher, P. G. (2017). The role of reactive oxygen species in the degradation of lignin derived dissolved organic matter. *Geochimica et Cosmochimica Acta*, 208:171–184.

Weiss, M. C., Sousa, F. L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., and Martin, W. F. (2016). The physiology and habitat of the last universal common ancestor. *Nature microbiology*, 1(9):1–8.

Wershaw, R. (1993). Model for humus in soils and sediments. *Environmental Science & Technology*, 27(5):814–816.

Wershaw, R. L. (1994). Membrane-micelle model for humus in soils and sediments and its relation to humification. Technical report, US Geological Survey; USGPO.

Wigner, E. P. (1950). The limits of science. *Proceedings of the American Philosophical Society*, 94(5):422–427.

Wu, L. M., Zhou, C. H., Keeling, J., Tong, D. S., and Yu, W. H. (2012). Towards an understanding of the role of clay minerals in crude oil formation, migration and accumulation. *Earth-Science Reviews*, 115(4):373–386.

Yang, C. (2016). The future of physics—revisited. *60 Years of Yang–Mills Gauge Field Theories: CN Yang's Contributions to Physics*, pages 1–10.

Yang, Z. (2014). *Molecular evolution: a statistical approach*. Oxford University Press.

Zaouri, N. A. (2013). Adsorption of different fractions of organic matter on the surface of metal oxide. Master's thesis, King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia.

Zhang, K., Zhu, X., Wood, R. A., Shi, Y., Gao, Z., and Poulton, S. W. (2018). Oxygenation of the mesoproterozoic ocean and the evolution of complex eukaryotes. *Nature Geoscience*, 11(5):345–350.

Zhang, L., Luo, L., and Zhang, S. (2012). Integrated investigations on the adsorption mechanisms of fulvic and humic acids on three clay minerals. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 406:84–90.

Zhang, S., Wang, X., Wang, H., Bjerrum, C. J., Hammarlund, E. U., Costa, M. M., Connelly, J. N., Zhang, B., Su, J., and Canfield, D. E. (2016). Sufficient oxygen for animal respiration 1,400 million years ago. *Proceedings of the National Academy of Sciences*, 113(7):1731–1736.