論文 / 著書情報
Article / Book Information

題目(和文)	鉄に富む温泉環境における微生物生態の地球化学的変動要因：メタゲノム解析と機械学習アプローチ
Title(English)	Investigating Geochemical Drivers of Microbial Ecology in Iron-Rich Hot Springs: Metagenomic and Machine Learning Approaches
著者(和文)	Li Hau Fatima
Author(English)	Fatima Li Hau
出典(和文)	学位:博士(理学), 学位授与機関:東京科学大学, 報告番号:甲第217号, 授与年月日:2025年3月26日, 学位の種別:課程博士, 審査員:上野 雄一郎,横山 哲也,尾﨑 和海,MC GLYNN SHAWN ERIN,癸生川 陽子,Gilbert Alexis
Citation(English)	Degree:Doctor (Science), Conferring organization: Institute of Science Tokyo, Report number:甲第217号, Conferred date:2025/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	要約
Type(English)	Outline

Investigating Geochemical Drivers of Microbial Ecology in Iron-Rich Hot Springs: Metagenomic and Machine Learning Approaches

A DISSERTATION PRESENTED
BY
FATIMA LI HAU
TO
THE DEPARTMENT OF EARTH AND PLANETARY SCIENCES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF SCIENCE
IN THE SUBJECT OF
EARTH AND PLANETARY SCIENCES

TOKYO INSTITUTE OF TECHNOLOGY
MEGURO-KU, TOKYO

June 2024

Chief Examiner

Ueno Yuichiro

Thesis Committee Members

Gilbert Alexis
Yokoyama Tetsuya
Ozaki Kazumi
&
McGlynn Shawn

Chief Examiner: Yuichiro Ueno Fatima Li Hau

Investigating Geochemical Drivers of Microbial Ecology in Iron-Rich Hot Springs: Metagenomic and Machine Learning Approaches

Abstract

Iron rich hot springs could represent analogue environments to early Earth ' s ferruginous oceans. To investigate the geochemical drivers of microbial community structure and metabolic potential I performed read-based, and metagenome assembled genome (MAG) analyses on sediment samples from five iron-rich hot springs of Japan. The study revealed the reductive pentose phosphate cycle as the most abundant carbon fixation pathway, widespread iron oxidation-reduction, and microbial groups showing high metabolic plasticity involved in the nitrogen and sulfur cycles. Temperature and pH were important drivers of variation. Reconstructed MAGs belonged to 63 previously unreported genera, revealing underexplored taxonomic novelty. A machine learning approach to test the feasibility of using microbial and pathway abundances to predict geochemical parameters in hot springs revealed that only some parameters, can be predicted with accuracy, and that thermophilic microorganisms are important prediction features.

Summary

The Precambrian Supereon, which includes the Hadean, Archean, and Proterozoic eons, represents nearly 85% of Earth's history (Stüeken et al., 2020) but is the period we know the least about due to a sparse rock record. While microorganisms were the dominant life forms on early Earth, much about their physiology and ecology remains unknown. The geochemical context in which organisms survived included an atmosphere with less than 0.1% oxygen compared to current levels and high carbon dioxide, along with oceans rich in dissolved ferrous iron (Catling, 2011; Holland, 1986; Lyons et al., 2014; Planavsky et al., 2011; Stüeken et al., 2020).Other factors such as temperature, pH, salinity, and nutrient availability remain uncertain (Crowe et al., 2014; Fakhraee Katsev, 2019; Krissansen-Totton et al., 2018; Watanabe et al., 2024) With the geochemical constrains we have on Precambrian oceans we can make some inferences on the availability of electron donors and acceptors microbes could have used for energy conservation, but availability does necessarily mean utilization, and function does not necessarily mean phylogeny. There are still many unknowns about the biosphere of the Precambrian, its connection with the environment and the possible ways we can interpret the geological record. One possible strategy to use the current knowledge on Precambrian environmental conditions and gain insights into the biosphere present at the time, is the use of analogue environment studies. Studied on analogue sites can help us elucidate the material cycling and microbial community structure of Precambrian-like environments and test our geochemical and microbial physiology models and hypotheses. Availability of environments in which soluble iron is present is limited, iron-rich hot springs represent unique opportunities to study microbial communities in such conditions. The focus of this study is CO2-buffered, naturally occurring, aerial, ferrous iron-rich hot springs. In this work, I use metagenomic techniques to analyze the metabolic potential, material cycling and species-level contribution of Bacteria and Archaea to the biogeochemical cycles in iron rich hot springs of Japan that represent analogues to Precambrian oceans. In Chapter 1, I summarize the current knowledge on Archean environmental conditions, the transition to the Proterozoic marked by the Great Oxidation Event (GOE) and putative roles of microbial communities to the material cycling on early Earth based on modern analogue environments. In Chapter 2, I discuss the metabolic potential of microbial communities in five iron-rich hot springs of Japan (Jinata, Okuokuhachikuro (OHK), Furutobe, Tsubakiyama, Kowakubi) studied with read-based metagenomic sequencing techniques and discuss the findings in the context of amplicon sequencing-based microbial abundance and geochemical conditions. The study reveals that the reductive pentose phosphate pathway is the main carbon fixation pathway in all springs, but performed by different organisms, namely Cyanobacteria, Gammaproteobacteria, and Zetaproteobacteria, no name a few; with the exception of

Kowakubi hot spring, in which the reductive Acetyl-CoA pathway dominated. Salinity and temperature were strong drivers of microbial community structure,whereas salinity, pH and temperature were the main drivers of metabolic potential, with highest temperatures showing more abundance of C-1 compound utilization pathways. Oxygenic photosynthesis and anoxygenic photosynthesis were negatively affected by higher dissolved organic and inorganic carbon as well as temperature, suggesting most phototrophic microorganisms are hindered by these conditions. In Chapter 3, I refine the metagenomic data to reconstruct metagenome assembled genomes (MAGs) to discern the species-specific contributions to the geochemical cycles. The reconstructed 242 MAGs belonged to diverse phyla and include 185 new species. The study revealed that most organisms are able to utilize several electron donors for energy conservation such as iron, sulfide, thiosulfate and organic carbon. In regard to iron, some organisms were able to perform both iron oxidation and reduction, and most of them encoded iron regulation genes, highlighting the importance of iron in the nutrient cycling. The Cyanobacteria MAGs revealed their potential for anoxygenic photosynthesis, further expanding their role in the springs. Several diazotrophs were found from several clades including the Cyanobacteria, Desulfobacterota, Firmicutes and Myxococcota, to name some. Metabolic plasticity is widespread. In Chapter 4, I describe the results of a feasibility study on the use of machine learning (ML) random forests to predict geochemical conditions in a data set of 925 hot springs from New Zealand, obtained from (Power et al., 2018). Microbial community structure at the species level and family level was compared to predicted metabolic pathways as input features on model accuracy. Species level data showed the highest accuracy amongst all datasets for predicting pH and Temperature with an accuracy of around 80% and 60%, respectively. Micronutrient concentrations were poorly predicted by either dataset. The present dissertation highlights hot springs as natural laboratories to test hypotheses on extreme environments on Early Earth, it allowed us to comprehensively understand the carbon, nitrogen, iron and sulfur cycles in Precambrian analogues and the contributions of species to them. Future work should focus on experimentally constraining the rates of assimilation of nutrients with the metabolic pathways uncovered in this study to obtain quantitative values of material cycling.

Contents

Listing of figures

Acknowledgments

Science is a collaborative endeavour,and the writing of this thesis could have not taken place without the support of many people. I want to specially thank my supervisors Shawn E. McGlynn and Yuichiro Ueno for all their support. To Nakagawa Mayuko, Luce M. Ward and Kakegawa Takeshi for their help since the beggining of the project. To Jeff Bowman for his help on the machine learning aspect of the project and for opening the doors of his laboratory. To the staff at the Earth-Life Science Institute for all their kindness and for reminding me to write my work reports on time. To the owners of the onsen facilities in Akita and Aomori (Wespa Tsubakiyama, Furutobe onsen, Kowakubi Hot Springs Shōhōen) for allowing us to take a piece of their precious springs with us. To Paula Prondzinsky for her friendship and scientific advice. This work was supported by JST SPRING, Grant Number JPMJSP2106.

A Jorge L. Suárez Caballero por ser una fuente constante de motivación, admiración, apoyo y dopamina. A Maria Elvira Li Hau por sus mitocondrias y por enseñarme a no ponerle limites a mis suenos.