PitA Enzymes: A Key to Early Life on Earth?

Hugues Sayer (hs252@st-andrews.ac.uk)
Supervised by Dr Joanne Boden
School of Earth and Environmental Sciences

1. INTRODUCTION

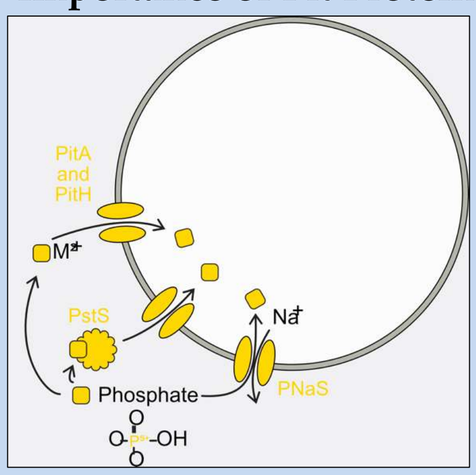
Life relies on six essential elements, including phosphorus. Early Earth's phosphorus was likely trapped in hard-to-access phosphate-metal complexes, posing a challenge for early life. This research project investigates how specific proteins, PitA and PitH which are phosphate-metal transporters, may have evolved to help early organisms access this important element, potentially solving the phosphorus scarcity problem during the Archean eon.

2. LITERATURE

The Archean

During the Archean eon (4.0-2.5 Ga), Earth's surface cooled to form continents and stable oceans, though the atmosphere remained anoxic (Kasting, 2014). This volatile period, marked by frequent volcanic eruptions and tectonic activity (Catling & Zahnle, 2020), saw the first signs of life, with microbial organisms like Bacteria and Archaea appearing in the fossil record.

Importance of Pit Proteins



Pit proteins transport inorganic phosphate into the cell of an organism. Some are always active (constitutively expressed), while others are regulated according to external inorganic phosphate (P_i) concentrations. This allows the organism to adapt to varying environmental P_i concentrations.

*This figure shows how a **cell** (black circle) uses **Pit proteins** (left yellow ovals) to gather **external inorganic phosphate** (yellow square) using **divalent metal cations** (yellow square labelled M*) (Boden et al. 2024).*

3. METHODOLOGY

Environmental & Taxonomic Distribution of PitA & PitH Homologs Using BLAST

For PitA & PitH to have ancient evolutionary origins, we need to find present day homologs across many environments and life forms. For PitA & PitH to be dated more accurately, we need to find in which organisms these homologs are found. To find and identify these homologs, searches were conducted on bioinformatics databases (NCBI & OceanGeneAtlas) to compile a comprehensive table for all homologs into Excel. All compiled environmental data was represented on a map. Only the taxonomic distribution of PitA homologs was covered using Krona (way to display taxonomic data) as PitH homologs had no HMM profile.

Producing a Phylogenetic Tree

Data on PitA Homologs was gathered from the bioinformatics databases and merged into a file. These homologs are proteins, composed of amino acid sequences which need to be aligned to produce an accurate phylogenetic tree. Using a remote computing cluster (type of supercomputer), these sequences were aligned and processed to create a phylogenetic tree using MAFFT and IQ-TREE commands. The output file generated was uploaded into TreeViewer to be edited.

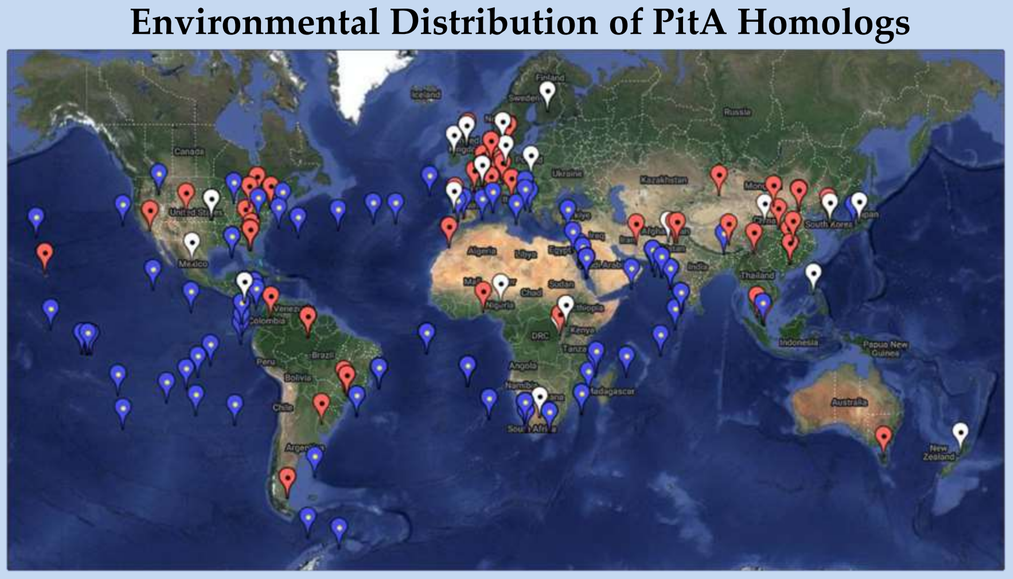
REFERENCES & ACKNOWLEDGEMENTS

- Boden, J.S. *et al.* (2024) 'Timing the evolution of phosphorus-cycling enzymes through geological time using phylogenomics', *Nature Communications*, 15(1). doi:10.1038/s41467-024-47914-0.
- Kasting, J.F. (2014) 'Atmospheric composition of Hadean–early Archean earth: The Importance of Co', *Earth's Early Atmosphere and Surface Environment* [Preprint]. doi:10.1130/2014.2504(04).
- Catling, D.C. and Zahnle, K.J. (2020) 'The Archean atmosphere', *Science Advances*, 6(9). doi:10.1126/sciadv.aax1420.
- **PitA Homologs – NCBI Page :** https://www.ncbi.nlm.nih.gov/genome/annotation_prok/evidence/NF033774/
- **PitH Homologs – NCBI Page :** https://www.ncbi.nlm.nih.gov/genome/annotation_prok/evidence/NBR010556/
- **OceanGeneAtlas :** https://tara-oceans.mio.osupytheas.fr/ocean-gene-atlas/

I would like to express my gratitude to the Laidlaw foundation for their support; for the opportunity I was given to expand my intellectual horizons alongside an exceptional group of peers.

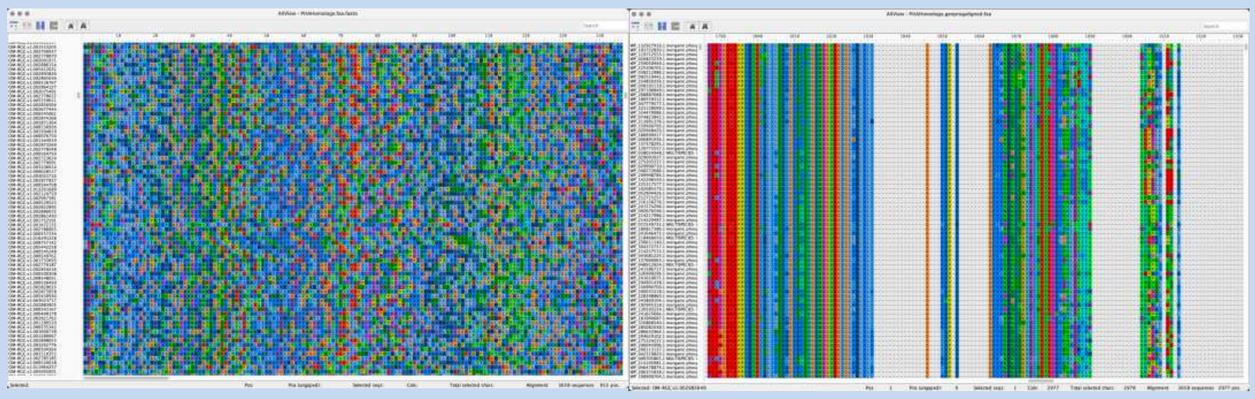
4. RESULTS

Environmental Distribution of PitA Homologs

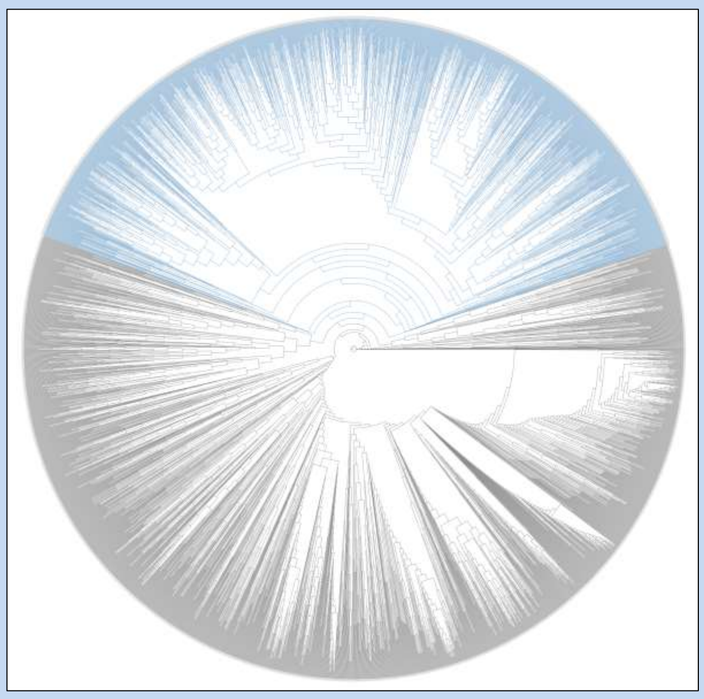


PitA was found in and around almost all oceans & continents. This supports the hypothesis that these proteins have ancient evolutionary origins. Krona also found that 81% of PitA homologs are found in *Pseudomonadota* which according to literature can be dated up to 3.1 Ga (Boden et al. 2024). This means PitA could have evolved anytime after 3.1 Ga, making it highly likely it emerged in the Archean.

Producing a Phylogenetic Tree



Result of the MAFFT alignment process



The marine sequences of the PitA homologs (blue) from Ocean Gene Atlas are clustered in one place on the tree. We can therefore conclude that the ability to use PitA to collect P_i from their environment was a single evolutionary event which happened in the Archean ocean.

5. CONCLUSION

1. Conducting environmental BLAST searches for PitA Homologs showed that these can be found in many different environments, confirming the idea that they have ancient evolutionary origins.
2. Conducting taxonomic BLAST searches for PitA Homologs showed that these are found primarily in the early bacterial lineage *Pseudomonadota* which can be dated to up to 3.1 Ga, making it possible for these proteins to have emerged during the Archean.
3. The evolutionary tree constructed further suggests that PitA originated from a single evolutionary event in a marine environment.

In conclusion, the evidence presented in this research project supports the notion that PitA was important in the survival and proliferation of early life forms in phosphorus-limited environments, showing a significant evolutionary innovation in the history of life on Earth.

LAIDLAW SCHOLARS

University of St Andrews