

Phylogenetic Oxygen Proxies: A. M. Caron and G. P. Fournier

Phylogenetic Proxies for the Rise of Atmospheric Oxygen. A. M. Caron¹ and G. P. Fournier^{1,1} Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge MA 02139

Introduction: The rise of atmospheric oxygenation on earth is widely agreed to have occurred in two major steps, the Great Oxygenation Event (GOE) around 2.33 Ga¹, and the Neoproterozoic Oxygenation Event (NOE) around 0.8 Ga², but oxygen levels between these two time-points remain highly unresolved. Mid-Proterozoic (1.8 to 0.8 Ga) oxygen levels have been estimated anywhere from 0.1%-40% of current atmospheric oxygen³, and proposed models of oxygenation over this time period include a steady post-GOE level, wild fluctuations, or an overshoot and subsequent crash scenario^(4,5). All previous work on the topic of atmospheric oxygenation has relied solely upon data from the geochemical record; we observe that an independent record should exist preserved within microbial lineages that experienced one or both of these events. Many genes found distributed across multiple microbial lineages encode proteins with functions dependent on the presence of molecular oxygen. These functions include preventing damage by oxygen free radicals (such as superoxides), catalyzing oxygen-dependent steps in biosynthesis pathways, oxygen sensing and stress responses, and the degradation of oxygen-dependent substrates. Reconstructing the evolutionary histories of these genes, we can interpret gene inheritance, horizontal gene transfer (HGT), and gene loss events as markers of periods of increased or relaxed oxygen selection pressure.

As lineages first experience increases in molecular oxygen levels, they will come under selective pressure to acquire genes to cope with or make use of this new environmental factor. If oxygenation continues, these genes will persist and tend to be vertically inherited among descendants as they undergo additional diversifications. A subsequent decrease in oxygen levels, or invasion of an anaerobic niche, may be accompanied by a loss of these same genes. Therefore, by mapping the HGT and gene loss events for oxygen-associated genes across multiple microbial lineages, one can build a clade-based profile of changing oxygen levels across both time and environments. Furthermore, if a calibrated molecular clock can be applied to these groups, the profile can be directly correlated with the geochemical record. This method allows for independent support for hypotheses regarding changing levels of atmospheric oxygen on Earth, with precision and on timescales relevant to the hypotheses surrounding the GOE and NOE events.

Our first analysis using a newly-designed bioinformatics pipeline including the Ranger II program⁶ for mapping oxygen-associated reticulate gene histories investigated the widely distributed Superoxide Dismutase (SOD) protein family, which catalyzes a critical step in detoxifying free radical oxygen⁷. We observe that while this gene has a broad phylogenetic distribution, it appears to lack a very ancient vertically inherited history in either Archaea or Bacteria, with only a few orders and classes of microbial lineages showing vertical inheritance of the gene (e.g., Methanosarcinales), implying molecular oxygen levels

increased before these groups diversified. Interestingly, the deepest branching lineage of Cyanobacteria, Gloeobacter, show an independent acquisition of the SOD gene via HGT, suggesting that this clade diverged before the GOE. Another interesting pattern is observed across the Crenarchaeota, in which several derived aerobic clades (e.g., Aeropyrum), all show independent HGT origins of their SOD genes. This supports a more recent origin for deep hydrothermal aerobes, consistent with hypotheses for a delayed, Neoproterozoic oxygenation of the deep ocean.

Future work using a fully automated pipeline incorporating multigene analyses and molecular clock models should permit a comprehensive, niche-specific mapping of the microbial genomic response to oxygen across planetary history. This will provide a novel proxy that complements traditional geological approaches, greatly improving our understanding of the alternative atmospheric regimes that dominated across different periods of geological time.

References:

- [1] Luo G. et al. (2016) *Science Advances*, 2.
- [1] Och L. M. and Shields-Zhou G. A. (2012) *Earth Science Reviews*, 110(1), 26-57.
- [3] Planavsky N. J. et al. (2014) *Science*, 346(6209), 635-638.
- [4] Canfield D. E. (2005) *Annu Rev Earth Planet Sci.*, 33, 1-36.
- [5] Lyons T. W. et al. (2014) *Nature*, 506(7488), 307-315.
- [6] Kordi M, and Bansal M. S. (2016) *ACM-BCB*, 297-306.
- [7] Zelko I. N. et al. (2002) *Free Radical Biology and Medicine*, 33(3), 337-349.