Network-level fossil of a phosphate-free biosphere. J.E. Goldford[1], H. Hartman[2], T.F. Smith[3], D. Segre[4], [1]Bioinformatics Program, Boston University, 44 Cummington Mall, Boston, MA 02215, e-mail: goldford@bu.edu, [2]Earth, Atmosphere and Planetary Science Department, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [3]Department of Biomedical Engineering and Graduate Program in Bioinformatics, Boston University, Boston, MA 02215, USA, [4]Department of Biology, Bioinformatics Program, Department of Biomedical Engineering, Department of Physics, Boston University, 44 Cummington Mall, Boston, MA 02215, e-mail: dsegre@bu.edu

Introduction: Phosphate is present in several key biomolecules, serving as an essential component of nucleotide triphosphates (ATP), coenzymes (NADH) and nucleic acids (RNA/DNA), used for biochemical energy transduction, redox biochemistry, and information storage, respectively. However, phosphate is geochemically scarce and poorly accessible, often serving as the limiting nutrient in a variety of modern ecosystems [1]. Currently, there is no consensus for a phosphate source in early life, athough recent work indicates that reduced phosphorus may have been available in Archaen oceans [2]. Even provided a source, the mechanisms of phosphate utilization and polymerization in early life remain unknown. These limitations have led to proposed models of early metabolic pathways devoid of phosphate, such as the "thioester world hypothesis" [3]. However, these models are typically applied to explain the prebiotic feasibility of generic reaction mechanisms. Could it be that, despite the ubiquity of phosphorus-containing metabolites in metabolism today, a substantial portion of intermediary metabolism existed prior to the incorporation of phosphate?

Addressing this question requires being able to efficiently model the complexity of large biochemical networks. Computational metabolic modeling enables the exploration of metabolic network organization at large scales. One method for global analysis of metabolism is the network expansion algorithm which simulates the generation of metabolic networks from a predefined set of seed compounds [4]. *In silico* networks are allowed to grow by first allowing all compounds to react based on a set of available reactions, forming products that can then be used in subsequent reactions. After a number of iterations, the simulated network stabilizes to a final set of reactions and metabolites (or scope) that are reachable from the initial seed set. Here we apply network expansion to the set of all known metabolic reactions (biosphere-level metabolism) to explore the capabilities of metabolic networks in absence of phosphate. In particular, we ask whether a metabolic subsystem could have in principle ensued from the expansion of a milieu of plausible prebiotic compounds lacking phosphate.

Results: Network expansion was performed using a simple set of prebiotically plausible molecules (Fig 1. top right box), generating a network with 260 metabolites and 315 reactions, enriched with reactions and metabolites involved in central metabolic pathways like the TCA cycle and amino acid metabolism.

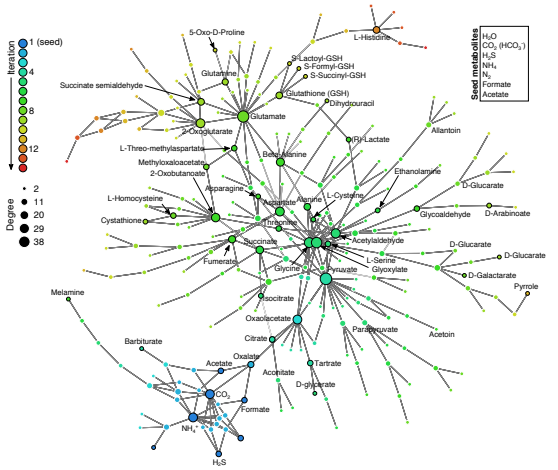

Figure 1: Core phosphate-free metabolic network

We find that features of enzymes in the network are overrepresented in proposed models of LUCA's proteome, and that enzymes in the network are enriched with iron-sulfur clusters, and metal coenzymes, corroborating previous models of ancient metabolic systems [5]. We then analyzed potential energetic constraints imposed by phosphate removal, and find that that thioesters may have been essential for overcoming local free energy barriers. We also find that inclusion of non-phosphate variants of modern day coenzymes may have enabled widespread network expansion without the use of phosphate. Our work corroborates previous proposals of a "thioester-world" and calls into question key assumptions surrounding the origin of living systems [6].

References:

[1] Schwartz, A. W. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361,** 1743–9 (2006) [2] Pasek M.A. *P.N.A.S.* (2008) [3] de Duve, C. (Neil Patterson Publishers, Carolina Biological Supply Company, 1991) [4] Ebenhöh, O *et al*. *Genome Inform.* **15,** 35–45 (2004) [5] Sousa, F.L. *et al*. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **368,** 20130088 (2013) [6] Goldford, J.E. *et al*. *Cell*, in press