Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Systems Biology

Modern views of ancient metabolic networks

Joshua E. Goldford[1] and Daniel Segrè[1,2,3,4]

Abstract

Metabolism is a molecular, cellular, ecological and planetary phenomenon, whose fundamental principles are likely at the heart of what makes living matter different from inanimate one. Systems biology approaches developed for the quantitative analysis of metabolism at multiple scales can help understand metabolism's ancient history. In this review, we highlight work that uses network-level approaches to shed light on key innovations in ancient life, including the emergence of proto-metabolic networks, collective autocatalysis and bioenergetics coupling. Recent experiments and computational analyses have revealed new aspects of this ancient history, paving the way for the use of large datasets to further improve our understanding of life's principles and abiogenesis.

Addresses

[1] Bioinformatics Program and Biological Design Center, Boston University, Boston, MA 02215, United States

[2] Department of Biology, Boston University, Boston MA 02215, United States

[3] Department of Biomedical Engineering, Boston University, Boston MA 02215, United States

[4] Department of Physics, Boston University, Boston MA 02215, United States

Corresponding author: Segrè, Daniel, Bioinformatics Program and Biological Design Center, Boston University, Boston, MA 02215, United States. (dsegre@bu.edu)

Current Opinion in Systems Biology 2018, ■:1–8

This review comes from a themed issue on **Regulatory and metabolic networks**

Edited by **Bas Teusink** and **Uwe Sauer**

For a complete overview see the Issue and the Editorial

Available online xxx

https://doi.org/10.1016/j.coisb.2018.01.004

Keywords

Metabolism, Evolution, Network expansion, Origin of life.

Introduction

The metabolic network of a cell transforms free energy and environmentally available molecules into more cells, moving electrons step by step along gradients in a complex energetic landscape [1,2]. The ability of a cell to efficiently and simultaneously manage hundreds of metabolic processes so as to accurately balance the production of its internal components constitutes a very complex resource allocation problem. In fact, it is only through recent systems biology research that we have begun to quantitatively assess this resource allocation problem at the whole-cell level [3,4]. A common perspective in the analysis of cellular self-reproduction is the notion that the genome, with its crucial information-storage role, is the central molecule of the cell, and that everything else can be collectively regarded as the machinery whose role is to produce a copy of the DNA. It is therefore not surprising that, as we struggle with the fascinating question of how life started on a lifeless planet, it is tempting to look for how a single information-containing molecule could arise spontaneously from prebiotic compounds. However, in spite of the appeal of thinking of DNA (or its historically older predecessor, RNA) as the central molecule who is being replicated in the cell, no molecule in the cell really self-replicates: the cell is a network of chemical transformations capable of collective autocatalytic self-reproduction. Collective autocatalysis is the capacity for a collection of chemicals to enhance or catalyze the synthesis or import of its own components, enabling a positive feedback mechanism that can lead to their sustained amplification. Combining this systems-level view of a cell with the argument of what is usually called the "metabolism first" view of the origin of life, one could propose that the ability of a chemical network to produce more of itself (or to grow autocatalytically) is and has always been a key hallmark of life [5–9]. An interesting modern version of this very same principle is embedded in one of the most popular systems biology approaches for the study of whole cell metabolism: this approach, based on reaction network stoichiometry and efficient constraint-based optimization algorithms, is commonly known as flux balance analysis (FBA) [3]. FBA solves mathematically the resource allocation problem that every living cell needs to solve in real life in order to transform available nutrients into the macromolecular building blocks that are necessary for maintenance and reproduction. When an FBA calculation estimates the maximal growth capacity of a cell, it essentially computes the set of reaction network fluxes that enable optimally efficient autocatalytic self-reproduction. While in cellular life this process is finely regulated and controlled, ancient life must have gone through many different stages of similar, but much less organized collectively autocatalytic processes. Thus, one of the key problems of the origin of life is the question of how an initially random path in the space of possible chemical transformations driven far from thermodynamic equilibrium could have ended up being dynamically "trapped" in a collectively autocatalytic state.

The focus on cellular self-reproduction as the fundamental level at which life and its origin should be understood is however too narrow. An exciting recent

development in systems biology of metabolism is the rise of methods to extend FBA models from the genome scale to the ecosystem level [10–12]. In addition to solving the resource allocation problem of metabolism for individual organisms in a given environment, these approaches take into account the fact that metabolites can be exchanged across species, giving rise to metabolically-driven ecological networks [11]. These advances suggest that metabolism may be best understood as an ecosystem-level phenomenon (Figure 1b), where the collective biochemical capabilities of multiple co-existing organisms may reflect – better than any individual metabolic network – an optimal capacity of life to utilize resources present in a given environment [9,13]. The ecosystem-level nature of metabolism is another feature of present-day life whose roots likely date back to the early stages of life on our planet. For example, the chemical networks that gradually gave rise to reproducing protocells may have wandered for quite some time in a broader chemical space, effectively generating molecular ecosystems before the rise of spatially and chemically well-defined cellular structures (see Figure 2).

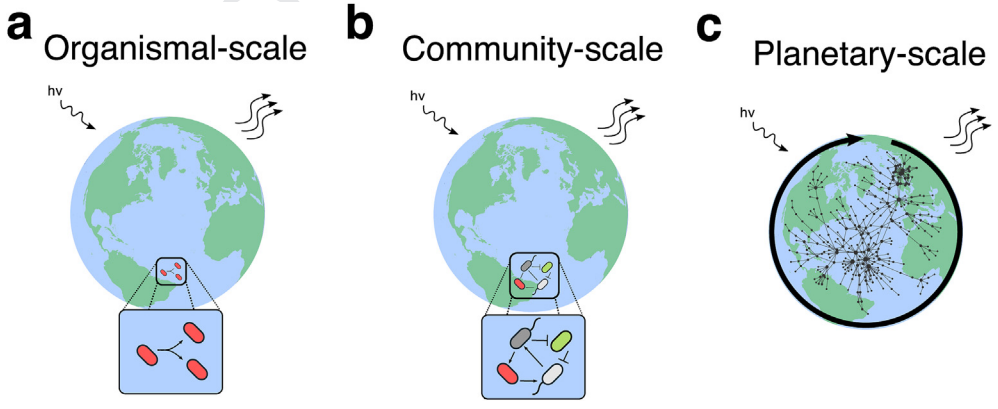
At an even larger scale, metabolism could be viewed as operating not just at the level of individual cells or ecosystems, but even as a planetary phenomenon, in which cellular processes collectively affect (and are affected by) the flow of molecules at geological scales (Figure 1c). The strong coupling between the metabolic processes of ecosystems and planetary-scale geochemistry [14,15] suggest that biosphere-level metabolism should be viewed as one of the natural scales for the study of life's history. A paramount challenge in the study of life's history is thus bridging the gap between material and energy fluxes at the biosphere scale, and detailed molecular mechanisms responsible for the properties of life at the cellular and subcellular level [16]. Bridging this gap could greatly benefit from the use of integrative models similar to the ones used in systems biology research and data science. In this perspective, we will discuss some recent system-level approaches that have provided new important insight into life's ancient history at multiple scales, highlighting the fact the metabolism and its multiscale nature – from the single reaction to the biosphere – are taking a center stage role in this endeavor.

Protometabolism before enzymes

A top-down reconstruction of ancient metabolic networks can be achieved based on the inferred history of gene families, using traditional phylogenomic techniques [17–20]. Leveraging information on the newly mapped genomic diversity of modern life [21], Martin and colleagues recently proposed a comprehensive phylogenetic reconstruction of the metabolic capabilities of the last universal common ancestor (LUCA), suggesting that LUCA was an autotrophic, thermophilic, N_2-fixing anaerobic prokaryote, living in hydrothermal vents and equipped with life's most complex molecular machines (e.g. ATP synthase) [22]. Although the details of LUCA's specific repertoire of metabolic enzymes are still subject of debate [23,24], these results corroborate the notion that LUCA was very complex, highlighting a massive gap in knowledge with regard to the transition from prebiotic geochemical processes to the biochemical complexity of LUCA and its progeny.

A major challenge in the study of the origin of metabolic networks is to gain insight on the structure of metabolic networks before LUCA and before the rise of genetic coding. At the core of this challenge is the question of whether and how metabolic reactions which depend on genome-encoded enzymes in modern cells could have

Figure 1



Metabolism at different scales. (a) Metabolic networks can be modeled at the organismal level, where environmentally-supplied resources, under an energy flow, are collectively transformed into biomass of the self-reproducing and evolving organisms. (b) At higher scales, metabolic networks can be viewed as an ecosystem-level phenomenon, where biochemical processes include metabolic exchange and competition between species. (c) Metabolism can be also considered a planetary scale phenomenon, whereby the free energy flow maintains global biogeochemical cycles.