Constraints that Deconstrain: Characterizing the Biological Hierarchy Across Levels of Organization. H. Kim[1], H.B. Smith[2], J. Raymond[2], and S.I.Walker[1,2,3] [1]Beyond Center for Fundamental Concepts in Science, Arizona State University, Tempe AZ USA [2]School of Earth and Space Exploration, Arizona State University, Tempe AZ USA, [3]Blue Marble Space Institute of Science, Seattle WA USA. Email: sara.i.walker@asu.edu

Introduction: One of the most characteristic properties of living systems is their hierarchical organization. Living structures are organized into multiple, nested structural levels ranging from the chemical reaction networks within cells, to cellular individuals, tissues, societies and ecosystems. Characterizing the structure of this hierarchy is widely regarded as one of the most important open questions in biology, necessary for solving the challenging open question of how such levels can emerge in the first place [1]. Yet, the question of how nesting structures within a hierarchy constrains the observed properties at each level or organization has not been addressed. Here we show that biochemical networks are organized to produce consistent topological properties across hierarchical levels, which imposes constraints on the properties of individuals at the lowest level. Our results demonstrate that by characterizing the structure of levels of organization across different scales, we obtain new insights into universal features of biochemical organization that are impossible to obtain by studying only one level in isolation.

Methodology:

Our dataset includes annotated genomes of 21,213 bacterial taxa and 730 archaeal taxa from the Pathosystems Resource Integration Center (PATRIC) [2], and 26 metagenomes sampled from hot spring ecosystems in Yellowstone National Park, each representing a microbial community defined by its geospatial location. Our motivation for selecting metagenomes sampled from Yellowstone hot springs for this study is that are among the most metabolically diverse ecosystems on the planet [3]. The samples in our dataset include communities living in temperatures ranging from 35-93 degrees C and pH from 2.1-8.3, and include both chemotrophic and phototrophic communities. From our dataset of annotated genomes and metagenomes we constructed individual and communal biochemical networks using reaction data cataloged in the Kyoto Encyclopedia of Genes and Genomes (KEGG) [4]. Each network is constructed from every catalyzed reaction (regardless of pathway) coded by the respective genome or metagenome, provided the reaction is catalogued in KEGG. For generating graph theoretic representations of these networks, we followed standard methods as developed in [5], where nodes are substrates and they are connected if an enzyme-catalyzed reaction transforms one substrate to another. These network representations allow systematic quantification of the topological properties of biochemical networks using the tools of graph theory and statistical mechanics, which we apply here to compare different scales of organization. We utilize the methods developed in [6] for determining the degree distribution. We also introduce two different types of controls on community-level network architecture to gain a better understanding of the reported trends. In the first, we randomly sample genomes to construct communities that are not co-evolved in the same geospatial location. In the second, we construct communities by randomly sampling artificial genomes that do not share a common biochemistry across species.

Results

From our genomic and metagenomic data, we find that the biochemical networks of individuals, communities and the entire biosphere display topological features that are universal and close to scale-invariant across levels organization. While many topological properties included in our study do appear scale-invariant, regardless of scale, the vast majority of networks in our dataset do not exhibit degree distributions consistent with a power law as previously claimed [5]. Artificial networks with the same exact topology as real individuals, but differing in their distributions of reactions shared between individuals, do not exhibit similar scaling behavior, indicating tight constraints on the relative composition of individuals in real-world systems.

Conclusion

By analyzing biological networks across hierarchical levels we can gain better insights into the structure of lower levels. Our results highlight the role of analyzing multiple, nested levels of organization in tandem to explain universal properties of life.

References:

[1] Szathmáry, E, and Maynard Smith, J.. (1995) *Nature* 374: 227-232.
[2] https://www.patricbrc.org/
[3] Swingley, W.D., et al. *PLoS One* (2012) 7: e38108.
[4] Kanehisa, M., and Susumu G. (2000) *Nucleic acids research* 28: 27-30.
[5] Jeong, H, et al. *Nature* (2000) 407: 651-654.
[6] Clauset, A., Shalizi, C. and Newman, M. (2009) *SIAM review* 51: 661-703.