Persistence Strategies in Biomolecular Network Architecture

Aishwarya Raj, Jay E. Mittenthal, Liudmila S. Mainzer

Abstract

The general framework of persistence strategies postulates that persistence in biological systems is achieved via a tradeoff of traits characterized by economy, flexibility or robustness. Here we investigate how these trade-offs can be quantified. We hypothesized that the structure and dynamics of biomolecular networks could differentiate between organisms of differing economy, flexibility, and robustness and subsequently classify unknown, newly discovered, or modified organisms in terms of their persistence trade-offs. Our approach is two-fold. First, we explored the network properties in theoretical terms, focusing on the simplest networks at first, then making predictions for more complex ones. Second, we used protein network data from real organisms to compare their size, centrality measures, extent of embedded redundancy, and node degree distribution. This ongoing exploratory work is aimed to provide support for the theory of persistence strategies and by generating testable hypotheses about molecular properties and network organization of living systems.

Introduction to Economy, Flexibility, Robustness (EFR)

Organisms are exposed to a variety of signals and stimuli. Responses of an organism to the signals of the umwelt help modulate the environmental effects on the organism's function, making the organism more **flexible**. Because processing signals is costly, organisms perceive and respond only to a small fraction of those signals. Organisms evolve properties of **robustness**, which allow them to continue functioning despite possible effects of the signals they do not perceive or process. The costs associated with flexibility and robustness are offset by the organism's budget of matter and energy. Flexible responses and robustness properties compete for this budget and are thus in a trade-off relationship, resulting in evolution of a particular **economy** – a method of meeting the organism's budget. Lineages evolve unique trade-offs among economy, flexibility and robustness (Figure 1). The coevolving economy/flexibility/robustness trio (**EFR**) is thus a dynamic attribute of every lineage, describing its particular **strategy of persistence**.

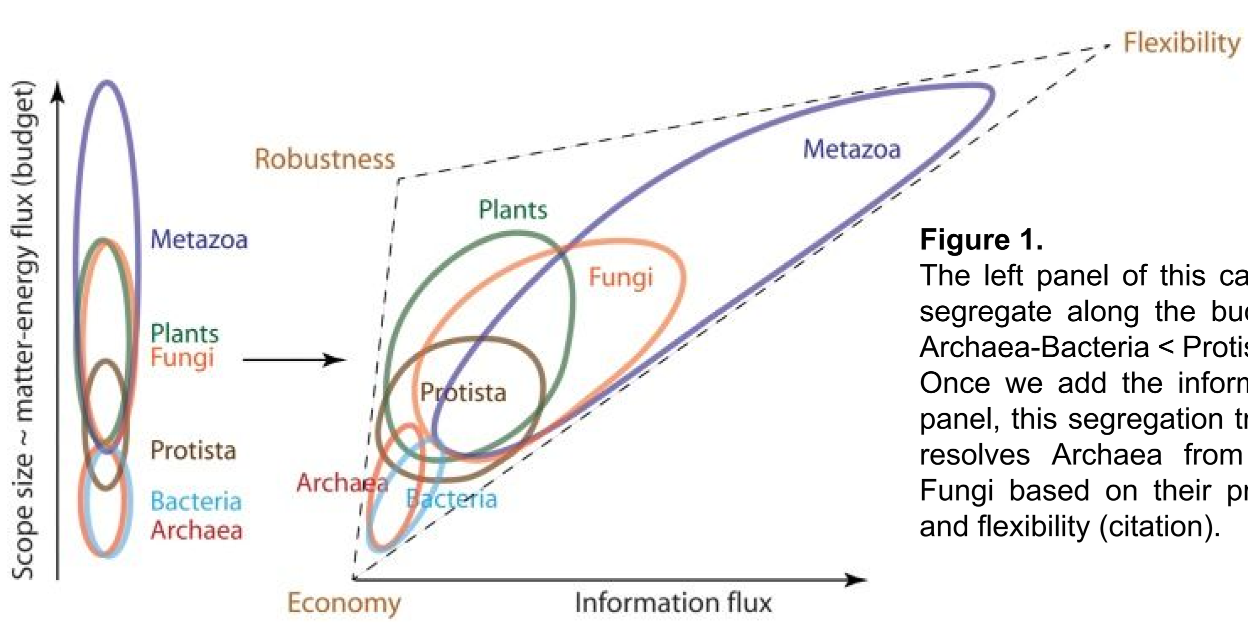


Figure 1.
The left panel of this cartoon shows that organisms segregate along the budgetary axis in the order of Archaea-Bacteria < Protista < Plants-Fungi < Metazoa Once we add the information flux axis on the right panel, this segregation transforms into a triangle that resolves Archaea from Bacteria and Plants from Fungi based on their propensity toward robustness and flexibility (citation).

Applying Network Theory to support EFR

A network consists of nodes (e.g. genes, proteins, neurons etc.), and their interactions are represented as the edges connecting the nodes. Edges can be undirected and denoted as lines, or have directionality attributes such as arrows; the directed edges may be excitatory or inhibitory (Figure 2, left). Structural properties of networks include degree (k), degree distribution(P(k)) and centrality [citation]. These properties describe patterns of node connectivity. For example, if node A has degree of k=3, node A would be connected to 3 other nodes (Figure 2, right) (citation). Individual node connectivity could facilitate integration or redistribution of information (Figure 3)

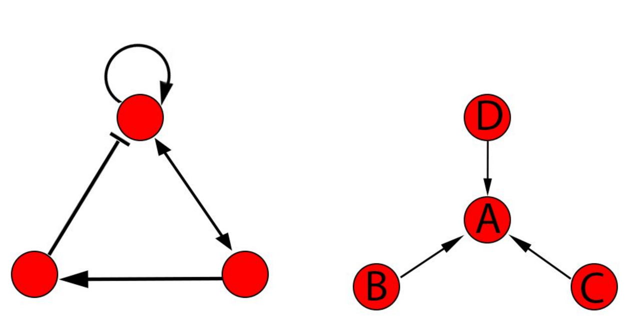


Figure 2.
Possible kinds of connections among nodes in a network: arrows indicate excitation, rectilinear ends indicate inhibition. Self-activation and self-inhibition are also possible. Degree of a node is equal to the number of connections it has to other nodes, whether those connections are excitatory, inhibitory, incoming or outgoing. Here node A has degree of 3, and nodes B, C, D have degree of 1.

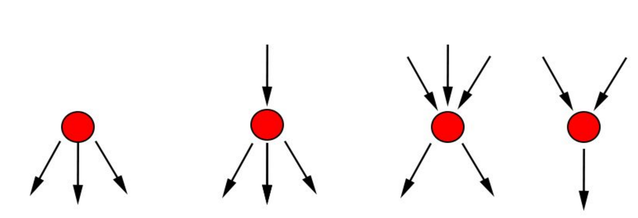
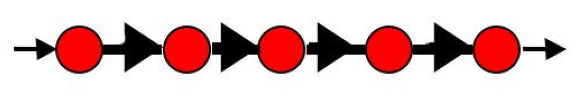


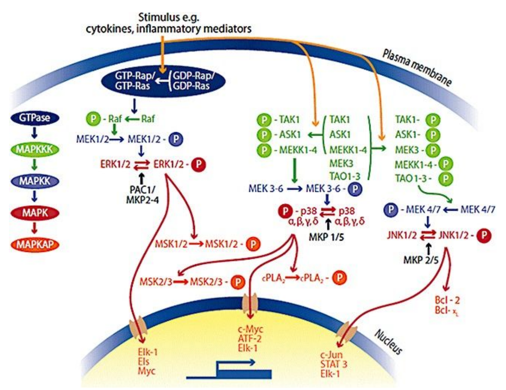
Figure 3.
Nodes with more outputs than inputs tend to distribute incoming information across the downstream network, potentially activating many effectors in response to few stimuli. Modes with more inputs than outputs tend to integrate the incoming information, potentially activating a single coherent behavior in response to a complex set of stimuli.

Interpreting EFR in terms of network architecture

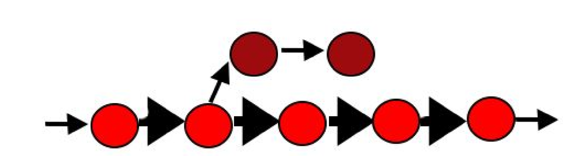
Economy: Economic networks are frugal in structure and function as to conserve matter-energy. We suggest that networks conferring economy have relatively few nodes that are arranged in sequence for faster operation. This minimizes redundancy in network architecture, reducing robustness. Unbranched structure leads to a single input/single output pattern, reducing flexibility.

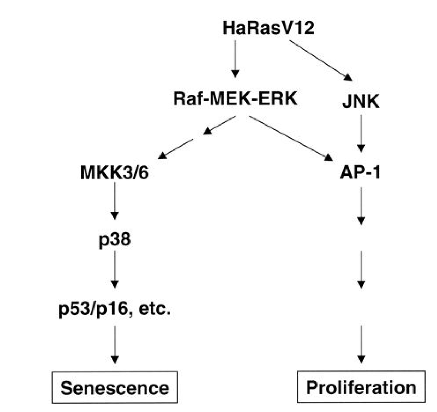


Sequential activation of the mitogen activated protein kinase family (MAPK) signalling cascade leading to protein and gene activation [7]



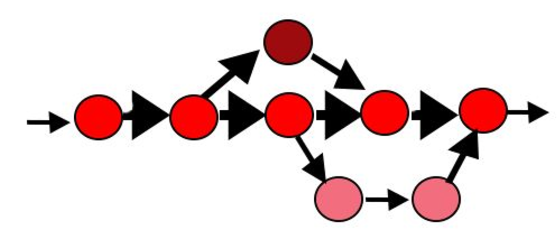
Flexibility: Flexible networks are defined as being able to respond to a greater number of stimuli with a greater number of outputs (downstream effectors). A node that has only one output can have only one response regardless of the stimuli. Similarly, a node that has only one input can only respond to one perceived signal regardless of the responses. However, a node with multiple inputs and/or outputs has the flexibility to respond to a number of stimuli with as many or more reactions. Based on the number of combinations of potential input/output stimuli for a given network, some network architectures may be classified as relatively more flexible than others.



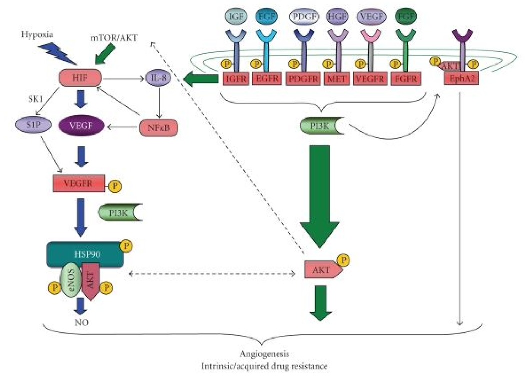


Schematic representation of a model indicating the role of p38 in oncogenic *ras*-induced premature senescence in primary fibroblasts[8]

Robustness: A robust network displays characteristics of redundancy, resulting in the ability to withstand numerous types of potentially environmental stimuli without failure. If one of the redundant nodes is damaged, the other redundant copies can still fulfill its function. Additionally, robustness can be conferred via an elevated number of distributors (kout>0 and Kin=0 or 1; Sergei and Maslov, 2005). This, however, may result in inefficiencies, i.e. longer time taken by a signal to traverse a network due to multiple paths that can be taken.



Hsp90-dependent modulation of proangiogenic signaling pathways in cancer. Hsp90 regulates multiple arms of angiogenic signaling in cancer[9].



a) Robustness can be quantified in structural terms, i.e. when a NW has a hub with many input and outputs, therefore conferring high flexibility. Eliminating that node/hub through damage results in a significant loss of flexibility. Thus by definition such structure is not robust.

b) Robustness can also be quantified in operational terms: symmetric and well-connected networks possess redundant paths for information flow.

All three characteristics represent types of organismal persistence strategies used for evolution and survival and consequently may be viewed as a trade off among each other: if a network loses one attribute, it may gain another.

Theoretical exploration of three-node network motifs

To explore how evolution of EFR can be supported via patterns of node connectivity, we modeled the simplest case of a three-node network. Three-node motifs have already received significant attention in the literature [5&6], thus providing us with ample ground to build on. We wrote a Python script (https://github.com/Araj6/EFR) that automates exhaustive enumeration of all possible three-node network motifs (total of 19,683; Figure 7). We assume that the top node receives the input signal, and eliminate all networks where a node is not connected to any other nodes. Many of the resultant motifs turn out to be redundant due to network symmetry. The remaining motifs can be grouped in terms of their likely dynamic patterns of output: always on, on-off, and oscillators with various period.

```
for first_loop in ['none', 'arrow_loop', 'inhib_loop']:
    for second_loop in ['none', 'arrow_loop', 'inhib_loop']:
        for third_loop in ['none', 'arrow_loop', 'inhib_loop']:
            for first_edge in ['none', 'arrow_none', 'none_arrow', 'inhib_none', 'none_inhib', 'arrow_inhib', 'inhib_arrow', 'arrow_arrow', 'inhib_inhib']:
                for second_edge in ['none', 'arrow_none', 'none_arrow', 'inhib_none', 'none_inhib', 'arrow_inhib', 'inhib_arrow', 'arrow_arrow', 'inhib_inhib']:
                    for third_edge in ['none', 'arrow_none', 'none_arrow', 'inhib_none', 'none_inhib', 'arrow_inhib', 'inhib_arrow', 'arrow_arrow', 'inhib_inhib']:
```

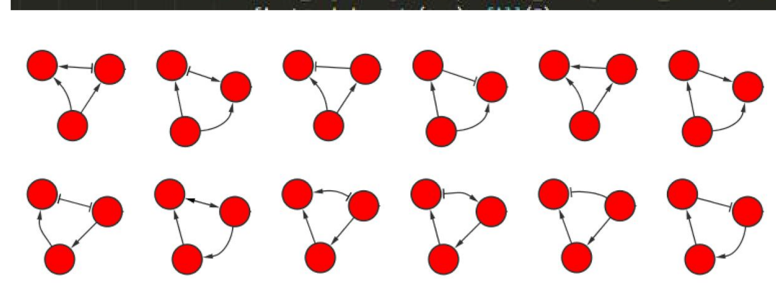


Figure 7.
Described above, part of the python script that was used to denote all possible directed interactions for 3 node subgraphs visualized in Cytoscape. Left, selected figures from the entire set of 19,683 possible directed 3 node graphs.

Next steps: evolvability of EFR

Our next step is exploring how the network motifs can be transformed one into another using a genetic algorithm, by modifying one element at a time (node, edge, strength of connection). We predict that there will be certain motifs that can be arrived at from many different starting points, whereas others best serve as starting points themselves. Additionally, small modifications to network motifs can result in a shift of EFR tradeoff from one persistence strategy to another. For example, an addition of a signal or a node can reduce economy but result in greater flexibility. Finally, by exploring the dynamics of each thri-node motif we can quantify the extent of E, F and R in it, and thus compare them. This process will serve as a foundation for comparing real organisms based on their biomolecular networks.

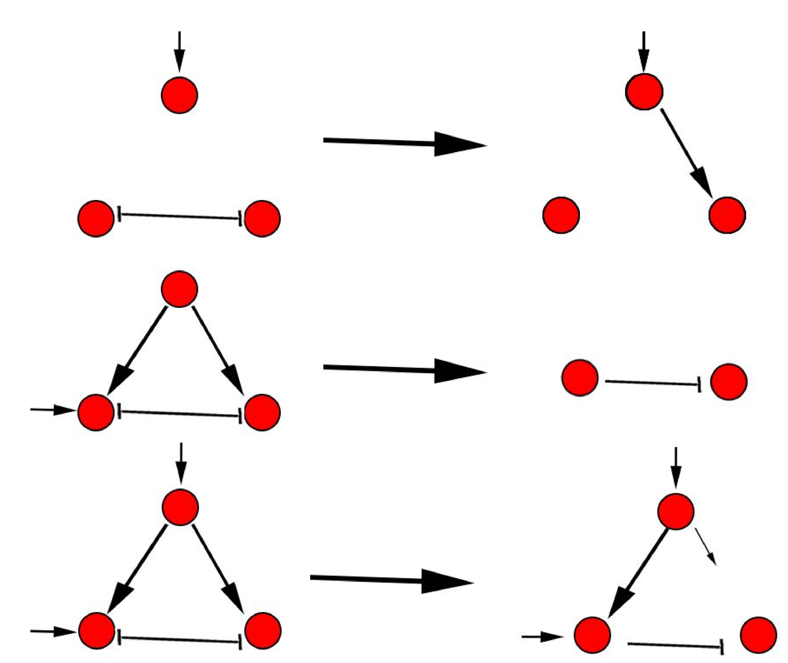


Figure 8.
Dynamic network behavior can be manipulated based on location, type and number of input stimuli. This type of manipulation can decrease economy and increase flexibility. The consequence of dynamic change in terms of type and number of stimuli can lead to transformation from persistence strategy to another.

References

1. Barabasi, A. L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature reviews genetics*, *5*(2), 101-113.
2. The Power of Power Laws. (n.d.).Retrieved April 25, 2017, from http://edgeperspectives.typepad.com/edge_perspectives/2007/05/the_power_of_po.htm
3. Yafremava, L. S., Wielgos, M., Thomas,S., Nasir, A., Wang, M.,Mittenthal, J. E., & Caetano-Anollés, G. (2013). A General Framework of Persistence Strategies for Biological Systems Helps Explain Domains of Life. Frontiers in Genetics, 4, 16. http://doi.org/10.3389/fgene.2013.00016
4. Van Noort, V., Snel, B., & Huynen, M. A. (2004). The yeast coexpression network has a small-world, scale-free architecture and can be explained by a simple model. EMBO Reports, 5(3), 280–284.http://doi.org/10.1038/sj.embor.7400090
5. Barabasi, A. L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature reviews genetics*, *5*(2), 101-113.
6. Shen-Orr, S. S., Milo, R., Mangan, S., & Alon, U. (2002). Network motifs in the transcriptional regulation network of Escherichia coli. *Nature genetics*, *31*(1),64-68. [7] Clauset, A., Shalizi, C., Newman, M. E. (2009). Power-law distributions in empirical data. SIAM Review, 51, 661-703.https://arxiv.org/abs/0706.1062
7. Broom, Oliver & Widjaya, Budiman & Troelsen, J & Olsen, J & Nielsen, Ole. (2009). Mitogen activated protein kinases: A role in inflammatory bowel disease?. Clinical and experimental immunology. 158. 272-80. 10.1111/j.1365-2249.2009.04033.x.
8. Wang, W., Chen, J. X., Liao, R., Deng, Q., Zhou, J. J., Huang, S., & Sun, P. (2002). Sequential activation of the MEK-extracellular signal-regulated kinase and MKK3/6-p38 mitogen-activated protein kinase pathways mediates oncogenic ras-induced premature senescence. *Molecular and cellular biology*, *22*(10), 3389-3403.
9. E Bohonowych, J & Gopal, U & Isaacs, Jennifer. (2010). Hsp90 as a Gatekeeper of Tumor Angiogenesis: Clinical Promise and Potential Pitfalls. Journal of oncology. 2010. 412985. 10.1155/2010/412985.

Acknowledgments

I would like to thank Katherine Kendig and NCSA Genomics for this opportunity, as well as the Illinois Informatics Institute, for providing me the fellowship to pursue this Investigation. I would also like to thank Jacob Heldenbrand and Ryan Chui for technical support throughout this project as well as Noah Flynn for whose initial efforts paved way for this research.