The gene-function relationship in the metabolism of yeast and digital organisms

Philip Gerlee[1], Torbjorn Lundh[2], Bing Zhang[3] and Alexander Anderson[4]

[1]Niels Bohr Institute
[2]Chalmers and University of Gothenburg
[3]Vanderbilt University
[4]University of Dundee
gerlee@maths.dundee.ac.uk

Many natural and artificial systems form structures which can best be viewed as networks consisting of a set of nodes and links connecting the nodes. This perspective has been helpful in elucidating the organisation of a variety of systems ranging from power-grids and the internet to protein interactions, gene regulation and cell metabolism. Many of these networks exhibit a scale-free degree distribution and therefore deviate from the classical description of complex networks which predicts a Poisson degree distribution, which for degrees larger than the average degree scales as an exponential distribution.

We have studied the metabolic gene-function network in yeast and digital organisms from the artificial life platform Avida. The gene-function network is a bipartite network in which a link exists between a gene and a function (pathway) if that function depends on that gene, and can also be viewed as a decomposition of the more traditional functional gene networks, where two genes are linked if they share any function. We show that the gene-function network exhibits two distinct degree distributions: the gene degree distribution is scale-free while the pathway distribution is exponential. This is true for both yeast and digital organisms which suggests that this is a general property of evolving systems. One possible explanation for this structure is that in the network the genes acquire new links according to preferential attachment while the pathways receive new links independent of their degree.

This hypothesis was tested in Avida by tracking the evolution of the gene-function network in repeated simulations and measuring the rate of link attachment. Here the single commands takes the role of genes and the functions are the evolved boolean functions for which the organisms are rewarded. The results show indeed that the a gene is more likely to become involved in new functions (i.e. increase its degree) the more links it already has. The link attachment of the functions on the other hand occurred independent of the degree. In real cells it is known that gene duplication is the main mechanism by which new genes are created. If the two genes would retain similar functionality then we would expect pathways which involve many genes to increase their degree. This is contradicted by the exponential degree distribution and the observations in Avida and suggests that the rate of the gene divergence in yeast must be high. The duplication of pathways/functions could on the other hand explain the scale-free distribution of the genes, and this mechanism has already been observed in Avida. Measuring the overlap between different pathways in terms of the genes which constitute them, showed that this also is a likely mechanism in yeast evolution. In conclusion we have presented a new way of analysing the gene-function dependence which sheds new light on the evolution of genes and functionality, and suggests that function duplication could be an important mechanism in evolution.