Comparison of expression in adult human tissues of genes involved in mendelian disorders versus non disease genes reveals consistency with a scale-free network model

S.Bortoluzzi, C. Romualdi, A. Bisognin, F. d'Alessi and G. A. Danieli

Dipartimento di Biologia, Universita' di Padova

In our laboratory, a fully automated procedure was developed for data mining from UniGene, the largest available collection of human expressed genes, and reconstructing expression profiles of human tissues. The complete list of the genes expressed in each of the considered tissues is available on line as a catalogue of transcripts. For each gene, the estimated level of expression in the considered tissue is available, each gene being linked to the corresponding LOCUSLINK page and subsequently to other relevant databases. Only those tissues were considered, for which a sufficiently large number of ESTs from unbiased libraries was available. In total, we reconstructed 58 expression profiles with slightly less than 30,000 represented genes. All these data are available at the web site http://telethon.bio.unipd.it/bioinfo/HGXP/

Expression profiles of different adult human tissues showed striking similarities: in every tissue, a relatively small number of highly expressed genes resulted to account for a very large fraction of the total transcriptional activity. We focused our attention on 15 normal adult human tissues: the expression in these tissues of genes reported in the OMIM morbid map, whose mutations cause mendelian disorders, and of a set of about 10,000 known human genes not involved in disease phenotypes, was considered. The classification of disease-genes according to their average level of expression in all the considered tissues, compared with the sample of genes of the reference set, showed that they are significantly more expressed than expected; in addition, when comparing the expression of genes involved in autosomal dominant diseases with the reference sample, the difference is even more evident. The emerging general picture is that the differentiated state of human tissues is characterized by high expression of a limited number of genes and by low expression of many genes. Disease-genes belong in general to the restricted group of highly expressed genes and their dominant mutations produce relevant phenotypic changes.

The complex metabolic system underlying cellular and tissue functions can be viewed as a huge network whose nodes are proteins and links potential chemical interactions between them. The model of "scale-free" inhomogeneous networks was recently proposed for biological systems: connectivity is maintained by a few highly connected nodes, whose removal drastically alters the network topology, whereas errors involving the majority of nodes with small connectivity do not increase the network diameter, having no considerable impact on network topology. Our observations seem to fit into the frame of a typical "scale free" network, which is simultaneously tolerant to random errors and fragile against the removal of highly connected nodes. Because of the properties of the "scale-free" network model, the system of differentiated tissues would maintain its characteristics during time, thus producing a fairly stable structural and functional tissue phenotype. Only "attacks" to key genes, by altering either their expression level or the quality of their product, would be able to significantly modify the cellular and the tissue phenotype.