**FROM GENOMICS TO EPIGENOMICS: DATA TRANSFERABILITY ACROSS THE EVOLUTIONARY SPECTRUM.**

GEORGE L. GABOR MIKLOS

The Neurosciences Institute, 10640 John J. Hopkins Drive,San Diego, California, USA, 92121

Keywords: genome sequences, transgenic rescue, phenotypic analysis, functional genomics

At what biological *levels* are the data from single celled organisms akin to a Rosetta Stone for multicellular organisms? How large is the gap between robogenomics, proteomics, and database mining, and the realities of functional genomics in the context of whole organismal phenotypes, and in particular, human disease phenotypes? To examine some of these issues, we utilized *Drosophila* as the experimental organism, and *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Homo sapiens* as the major comparators. We characterized a saturation mutagenized 67kb region of the fly genome by gene deletions, transgenic rescues, phenotypic analyses, genomic and cDNA sequencing, reverse transcription-PCR studies, bioinformatic analyses and evolutionary comparisons. These multi-level data sets revealed 12 different proteins, most of which are absent from bacterial databases, half of which are missing from *Saccharomyces cerevisiae*, yet nearly all of which have relatives in this worm and in human beings. Furthermore, our deletion studies reveal that a morphological phenotype is seldom observed when these genes are removed from the genome. These data are little different to those from the much larger 2.5 megabase Adh region of the fly (Rubin and Ashburner, unpublished). They pinpoint the pragmatic bottlenecks in functional genomics, and reveal the emerging problems with data transferability above the levels of genes and proteins, particularly when comparisons are made across a broad evolutionary spectrum. At these higher levels the popular notion that the yeast genome is *the* biological Rosetta stone, (in which it is claimed that the meaning of the sequence of human disease genes is routinely deciphered using information from yeast and worms), has little predictive significance and is more hype than reality. However newer transgenically-based mis-expression screens in *Drosophila* and *Mus*, parallel analyses of genetic selections in yeast, and coherency pattern analyses of gene networks, all offer restricted insights into organismal function at different biological levels. We conclude that the industrially scaled robogenomics approaches in model organisms (using chip-based transcriptome and proteome analyses combined with mass spectrometric measurements), together with appropriate network analyses, may have great impact on our understanding of a limited subset of non-neuropsychiatric human diseases. This will be made more likely if the transgenic and high throughput technologies can be realistically linked to epigenetic and phenotypic analyses of human diseases, particularly if these diseases are evaluated in the context of different genetic backgrounds.

**References:**

1. Miklos and Rubin, 1996. Cell, 86, 521-529;
2. Maleszka et al., 1998. PNAS, 95, 3731-3736;
3. Wen et al.,1998. PNAS, 95, 334-339;
4. Rorth et al., 1998. DEVELOPMENT, 125, 1049-1057;
5. Cho et al., 1998. PNAS, 95, 3752-3757.