

Automated quantitative analysis of phenotype similarities

Eckley, D. M., Shamir, L., Macura, T., Orlov, N., Goldberg, I. G.,

The post-genomic era has revealed thousands of uncharacterized genes. Assigning their myriad functions requires a similarly systematic approach. Elucidating genetic pathways initially requires sorting genes into appropriate functional groups. We are approaching this task using machine vision algorithms to deduce gene function from terminal phenotypes caused by single-gene knockdown in cultured cells. Images of these phenotypes are acquired using an automated Deltavision microscope. The 3D images (1024x1024x20) are deconvolved, maximum-intensity projected, and split into 16 equally sized tiles. A multi-purpose image classifier (WND-CHARM) is applied to quantitatively measure the morphological similarities of the various resulting phenotypes. The computed similarities between phenotypes are visualized in the form of phylogenetic trees similar to those used when comparing genes by sequence similarity. The efficacy of the proposed method was evaluated experimentally by using a group of 19 genes with published phenotypes and one gene of unknown function (CHD1). Analysis of phenotype similarity revealed several functional classes, which appeared as clusters in the resulting dendrogram. Genes that were predicted to have similar phenotypes because of their enzyme/substrate relationships, or participation in the same protein complex had the closest association in the dendrogram, while genes that were predicted to have phenotypes unique in this gene subset such as CHD1 (unknown function), and Diap1 (Cell death), were isolated. We conclude that automated analysis of image similarity can be used in morphological screening of RNAi libraries to group genes by phenotype and thus characterize genes with unknown function.