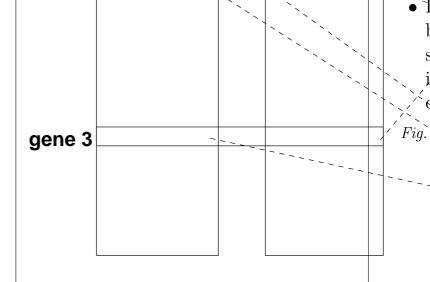
Functions of human genes are often studied indirectly, by model organisms such as the mouse. An underlying assumption is that genes with a common evolutionary origin i.e. orthologous genes, have similar functional roles in both species.

It is critical to know whether the function of orthologues is really similar. Orthologous gene groups with unexpectedly many similarly expressed gene pairs may highlight important physiological similarities between the species. Diverged gene function may refer to significant evolutionary changes.

Here a dependency exploration tool called associative clustering (AC) [2] is used to identify orthologous man-mouse gene pairs with potentially interesting regularity in their expression (Fig. 1).



Repeated AC (by bootstig gene 1 gene 3 gene 2 gene 5 gene 3 gene 7 gene 4 gene 9 gene 7

OVERLAPPING CLUSTERS

Principle of AC. Two sets of gene expression profiles are clustered simultaneously, and the results are represented on a contingency table where clusters of the two data sets correspond to the rows and columns of the table, respectively, and are called the margin clusters. Table cells, i.e. cross clusters contain orthologous gene pairs of the two data sets. Cluster centroids are optimized to produce a maximally dependent table, in the sense of the Bayes factor of dependent vs. independent margins. Potentially interesting co-occurrences are revealed by cross clusters with significant deviation from the hypothesis of independent margins. Reliability of the results is enhanced by bootstrap i.e. repeating the clustering with slightly differing samples; frequent co-occurrences are selected for further analyses from a co-occurrence tree with a specified threshold.

Methods

• AC explores dependencies in paired

Conti to one margi from

g

g

ge

ge

