

Poster A-25

Patterns of Gene Colocalization in the Human Genome



Authors:

Par Engstrom (*Computational Biology Unit, Bergen Center for Computational Science, University of Bergen, Bergen, N*)

David Fredman (*Computational Biology Unit, Bergen Center for Computational Science, University of Bergen, Bergen, N*)

Boris Lenhard (*Computational Biology Unit, Bergen Center for Computational Science, University of Bergen, Bergen, N*)

Short Abstract: We present a new strategy to explore gene order in annotated genomes. Our algorithm searches for clusters consisting of genes from different functional groups and reports recurring constellations observed more often than expected. We describe human gene clusters that are conserved in evolution and encode proteins with different regulatory functions.

Long Abstract:

Although it has been established that eukaryotic genes are nonrandomly ordered along chromosomes, we have only a partial understanding of the determinants of eukaryotic gene order. In addition to the occurrence of clusters of tandemly duplicated genes, several deviations from random gene order have been described. Genes with detectable expression in a broad range of tissues (housekeeping genes) have been shown to be clustered in Giemsa-negative or -light-staining regions in the human genome. Clusters of coexpressed genes have been reported in several eukaryotic genomes. Eukaryotic genes involved in the same metabolic pathway have also been shown to be less dispersed than expected if genes were randomly ordered.

We present a new strategy to explore gene order in annotated genomes. While previous work has focused on finding clusters of genes for which expression data or annotation suggests common functional properties, our algorithm searches for clusters consisting of genes from different functional groups. The algorithm reports recurring clustering patterns observed more often than expected by chance. The algorithm is designed to be unaffected by the occurrence of tandem duplicates and variations in gene density along chromosomes.

We have applied the algorithm to the human genome, using Ensembl gene models and Gene Ontology annotation. We describe the resulting clusters, which include groups of genes involved in regulation of gene expression at different levels. Using alignments between the human and mouse genomes, we further demonstrate that observed clusters tend to be maintained in evolution. Finally, we discuss possible explanations for the observed gene clustering, e.g. regulatory properties of the regions harboring the clustered genes.