Computational analysis of individual human genomes

Daslav Hranueli, *Faculty of Food Technology and Biotechnology*

Abstract
The genome sequences of a small number of individuals have already been determined. "New generation" DNA sequencing technology should lead to rapidly falling prices and the widespread determination of individual genome sequences. A major problem will be the provision of expert advice to interpret the results, both for the individual and their physician. An integrated set of computer programs (***ClustScan****)* was developed initially to annotate modular biosynthetic clusters in bacteria that are important for the production of antibiotics and other pharmaceutically important products. The concept of ***ClustScan*** is to use a top-down approach incorporating knowledge in an ordered way (using XML) to produce results in a form useful for Biologists including a special graphical user interface. The versatility of this program package has been demonstrated by annotating biochemical pathways in microbial, invertebrate animal and metagenomic datasets. The open architecture of ***ClustScan*** allows easy integration with other programs and updating to incorporate new knowledge. The analysis of an individual human genome requires a number of different sorts of analyses. This includes detection of alleles (both known and novel) of genes as well as determination of haplotypes that are linked to QTLs of medical and pharmacogenomic significance. The results of the analyses must be integrated with existing knowledge to produce a medically useful tool. The concepts used in ***ClustScan*** seem ideally suited to generate a novel tool for automatic individual genome annotation, ***GenTell*** (***Gen***ome In***Tell***igence). We would welcome collaboration with clinicians in the development of ***GenTell***.

Key words: Human genome sequencing, comparison via automatic annotation, ***GenTell***