

Bridging a Gap between DNA sequences and expression patterns of genes

Shinichi Morishita

Abstract

The completion of sequencing human genome would motivate us to map millions of human cDNAs onto the unique ruler "genome sequence", in order to identify the exact address of each cDNA together with its exons, its promoter region, and its alternative splicing patterns. The expression patterns of some cDNAs could therefore be associated with these precise gene addresses, which further accelerate studies on mining correlations between motifs of promoters and expressions of genes in tissues.

Towards the realization of this goal, we have developed a time-and-space efficient software named SQUALL that is able to map one cDNA sequence of length a few thousand onto a long genome sequence of length thirty million in a couple of minutes on average. Using SQUALL, we have mapped twenty thousand of our Bodymap (<http://bodymap.ims.u-tokyo.ac.jp>) cDNAs onto the genome sequences of Chr.21st and 22nd. In this talk, I will report the status of this ongoing project.

Curriculum Vitae

Shinichi Morishita

Shinichi Morishita is an associate professor of University of Tokyo. After graduating from Department of Information Science, University of Tokyo, he worked with IBM Tokyo Research Laboratory, Stanford University, IBM Almaden Research Center, and Institute of Medical Science, University of Tokyo. His major research interests are mathematical logic for computer science, algorithms, computational complexity, database system, implementation, data mining technology, and genome informatics.