[Talk / CBRC2012:The Impact of Next Generation Sequencers]

Sequence alignment methods useful at various stages of genome sequence analysis

Osamu Gotoh

Comparative Genomics Research Group Leader, CBRC, AIST

Sequence alignment is a fundamental tool in biological sequence analyses. In the present era of huge genomic and transcriptomic sequence data, the importance of alignment methods is ever increasing. We are developing several lines of alignment tools that are useful at various stages of genomic sequence analysis, including mapping of NGS-generated reads to reference genome, splice variant detection, and mutual alignment of gigantic whole genomic sequences. An on-going study is aimed at RNA-seq analysis with mid-to-long reads generated by the latest sequencers. Our algorithm can deal with these relatively long but error prone reads much better than conventional tools that are optimized for short errorless reads. Another topic will be on gene prediction based on profile-based spliced alignment in combination with gene-structure-aware multiple protein sequence alignment. Some results obtained from real data will be presented.