Bioinformatics for Functional Genomics in BRIC

Yang-Suk Kim, Sujin Choe, Ji-Young Oh, Sung-Whan Kim, Dong-Sun Park, Tae-Jin Ahn, Hong-Gil Nam

Biological Research Information Center, Pohang University of Science and Technology, San 31 Hyoja-dong, Nam-Gu, Pohang, Kyoungbuk, South Korea

BRIC, as a public bioinformatics center in South Korea, has been performing several activities for promoting research power of Korean biologist. As a starting point of bioinformatics, we have constructed the Korean sequence database, GeneNuri, to manage genetic resource in Korea. We have introduced object-oriented concept to express complex hierarchical data structures and to combine data set and application software tightly. We also developed web-based sequence submission system, Gene-In. It provides user-friendly interface and has automatic annotation such as verifying misspelling, sequence length, and correct translation. Gene-In also provides independent submission tool for users who want to submit large number of sequence, such as multi EST sets.

Another project for bioinformatics service is to develop the system for sequence meta-search engine. It interacts with several biological databases, such as GenBank, BLOCK and PDB and performs automatic search by user's choice. The searched results are automatically combined and provided to users. Another function of meta-search engine is scheduling. Often biologists want to search their sequence periodically to the updated database. To satisfy this request, meta-search engine performs search periodically. When search is done, meta-search engine compares the previous one with the new results and can provide users for only updated result.

For functional genomics, we are preparing an intelligent database management system. Object concepts will be applied to database for complex hierarchy. In addition to data storage, application programs for sequence analysis will be tightly linked to database. For database management system, we will introduce the concept of ruleDB, which will contain the information for the most optimal search condition and automatic annotation for functional prediction. The ruleDB will be constructed by machine-learning technologies.