First ClustalX-MPI for Ultra-fast Protein and Gene Discovery

Teaho Kim, Jin Han, Jae boum Youm, Nari Kim, Won sun Park, Sunghyun Kang, Dang Van Cuong, Hyoung kyu Kim, Tran Min Khoa, Vu Thi Thu, Hyunju Kim, Hyejin Moon, Hyunsuk Lee, Euiyong Kim and Hyun Joo*

Mitochondrial Signaling Laboratory, Department of Physiology and Biophysics, College of Medicine, Cardiovascular and Metabolic Disease Center, Biohealth Products Research Center, Inje University, Busan 614-735 Korea

Bio-Grid computing is an emerging area that utilizes geographically distributed commodity computers as a virtual supercomputer. With the growing amount of data, it becomes impossible to analyze massive gene or protein sequences manually. Based on Clustal algorithm, we firstly developed ClustalX-MPI tool for more convenient and powerful sequence searching, which allows user-friendly GUI interfaces during similar sequences or functions in the genome-wide sequence spaces, within billions of nucleotides. In an integrated scale, recently we designed a high performance distributed memory computation system to solve complex biological sequence-to-function problems. The machine was applied for implementing a LINUX-based high performance MSA algorithm, which aligns huge protein or genome sequences. These programs also need an compensation for various mutations in order to identify sequences that are closely related, but not identical.