Biocomputing resources at the DKFZ

Peter Ernst, Mechthild Falkenhahn, Karl-Heinz Glatting, Agnes Hotz-Wagenblatt, Clemens Suter-Crazzolara and Sandor Suhai
Department of Molecular Biophysics, DKFZ, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

c.suter-crazzolara@dkfz-heidelberg.de

With the advent of biomedical research, and particularly several genome projects, a wealth of sequence data is currently being collected. The speed of data collection exceeds the rate with which these data are analysed. The gap between these two will dramatically widen in the next few years. Therefore, researchers need rapid and straightforward access to software programs and databases, both to analyse new sequences and to obtain biological information.
The Department of Molecular Biophysics (DKFZ, Heidelberg) supports scientists involved in all aspects of biomedical research. Our services include:

- full access to databases, analysis tools, computers and network hardware for our users. Sequence analysis can be carried out with more than 200 programs, present in packages such as HUSAR/GCG, GDB/OMIM, STADEN and SRS. Additional software for the Bioccelerator allows extremely fast database searches. We offer more than 60 different databases[1] for use by the scientific community, including the complete EMBL and GENBANK nucleotide databases (updated daily), protein sequence databases (e.g. SwissProt and PIR, updated weekly) and more specialised collections such as the HIV-, TransFac- and PROSITE- databases, or the complete genomes of organisms such as yeast and E. coli. We offer extensive user support (bioinformatics courses and workshops, telephone hot-lines, newsletters).
- Our World Wide Web interface to GCG/HUSAR[2] allows user-friendly access to applications offered to the research community (*http://genome.dkfz-heidelberg.de*). This interface was developed in collaboration with the European Bioinformatics Institute (Hinxton, UK) and is being extended constantly.
- We are active participants in the German and European Genome Projects. For the latter we function as the European Data Resource (EDR). Furthermore, our institute maintains GENIUSnet, the German EMBnet node. EMBnet is a group of 28 nodes throughout Europe. The combined expertise of the nodes allows EMBnet to provide a service to the European molecular biology community which includes data and software distribution to users, independent research, networking and training.
- In participation with our customers, we develop specialised programs. One of these projects is EST-CLUSTERING, an automated process which identifies overlapping sequences and assembles them into one contig. Integration of existing programs into HUSAR (such as DIALIGN, CAP or DSC) is an additional pivotal part of our service.

[1]Ebeling, M., Suhai, S. ”Molecular databases on the Internet” J. Mol. Med. 75 (1997) 620-3. [2]Senger, M., Flores, T., Glatting, K., Ernst, P., Hotz-Wagenblatt, A., Suhai, S. “W2H: WWW interface to the GCG sequence analysis package” Bioinformatics 14 (1998) 452-7.