

# Bioinformatics at the ABCC

THE ADVANCED BIOMEDICAL COMPUTING CENTER (ABCC) IS A CENTER FOR RESEARCH AND SUPPORT IN COMPUTATIONAL BIOLOGY, GENOMICS, PROTEOMICS, BIOINFORMATICS AND MOLECULAR MODELING.

The ABCC employs a highly skilled informatics staff with concentrated expertise in proteomics and bioinformatics to ensure that the growing body of information from molecular biology and genome research is available to the scientific community in ways that promote scientific progress.

The staff helps users develop and utilize efficient high performance applications and make effective use of resources by supporting programming tools, applications, and libraries. Training, consultation and collaboration with researchers are available in group sessions and one-on-one. The ABCC is also strongly involved in research projects in bioinformatics and computational molecular biology.

## Databases and Bioinformatics Software

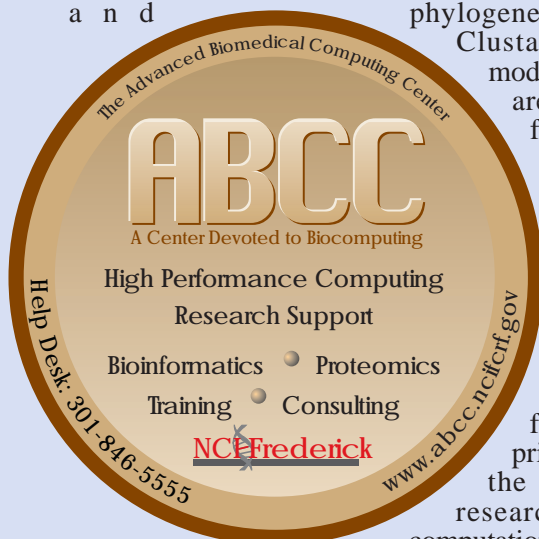
The ABCC informatics staff builds, maintains and supports program suites, databases and information services. The same programs and databases available at the national genome sequencing centers are available locally to ABCC users and the databases are distributed on the center's ftp site – <ftp://ftp.ncifcrf.gov>. Up-to-date, local copies of such databases as Genbank, PDB, PIR, SWISS-PROT, Genpept, NRDBNucl and NRDBProt provide ABCC users with faster and more reliable access. Large subsets from these databases are also created and distributed through the ftp facility for other large bioinformatics users such as pharmaceutical and biotech companies.

A comprehensive and diverse set of software resources in the field is available to researchers. The most popular program suites (and many others) used for assembly (Phred, Phrap, stacks), sequence analysis (GCG), linkage (Fastlink, RH-map, SimLink) and

phylogenetic analysis (Phylip, ClustalW), and molecular modeling (InsightII, Look) are accessible via user-friendly, web-based interfaces.

Sophisticated, customized processes have been designed and created for connecting multiple applications (e.g., procedure for micro-satellite sequence finding and subsequent primer design) to reduce the amount of time the researcher expends at the computational phase. Customized

processes have also been developed in-house to alter and enhance program functionality. Such is the case with the PCR primer prediction procedure to produce spread-sheet ready output.



## Training and Assistance

Education and user assistance have always been an integral part of the center's mission. The ABCC offers a variety of educational forums aimed at training scientists and researchers: workshops, classes, seminars, one-on-one discussions, a HELP DESK (301-846-5555), and a web-based FAQ database – <http://www.abcc.ncifcrf.gov/support.shtml>. Introductory and advanced classes are taught through "Genomics/Proteomics" and "Internet Tools" tracks. New courses such as "High Throughput Sequence Analysis Methods" and "Introduction to the Geospiza Sequence Storage System" have been added to complement the ever popular "Sequence Analysis at the ABCC" workshop. Training schedules, course descriptions, as well as online registration, are available at <http://www.abcc.ncifcrf.gov/training/index.shtml>

- *Introduction to Sequence Analysis at the ABCC*
- *High Throughput Sequence Analysis Methods*
- *Introduction to the Geospiza Sequence Storage System*
- ... and others.

## Accessing Resources

The ABCC's bioinformatics resources are available at no cost. Application forms can be obtained at [http://www.abcc.ncifcrf.gov/apps\\_apply.shtml](http://www.abcc.ncifcrf.gov/apps_apply.shtml) or by sending e-mail to [apps-request@ncifcrf.gov](mailto:apps-request@ncifcrf.gov). An account enables you to access a wide variety of runnable programs in proteomics, genomic analysis, and comparison and searching on the ABCC's web site at <http://www.abcc.ncifcrf.gov/app/htdocs/appdb/>. In addition, consultation and training are available from the ABCC staff on a variety of bioinformatics topics. As new applications and data sources are made available we are implementing web-based access methods wherever possible for these applications. Other applications are available only through Unix interfaces.

## Web-Enabled Applications at the ABCC (partial listing)

Applications	Description
<b>Proteomics</b>	
<a href="#">Protein Prospector</a>	Mass Spectrometry related proteomics searching tools.
<b>Comparison &amp; Searching</b>	
<a href="#">Blast</a>	Search Databases with Sequence Query
<a href="#">DocSearch</a>	Search Databases with Text Query
<a href="#">Fasta</a>	DNA Sequence Scanner
<a href="#">FindRepeats</a>	ABCC Repeat Finder Program
<b>Primer Selection</b>	
<a href="#">MSP</a>	MicroSatellite Target Identification Program
<b>Genomic Analysis</b>	
<a href="#">GenScan</a>	Scan Sequences for Exons/Signams
<a href="#">Repeat Masker</a>	Screen out Repeat Sequences
<a href="#">Screen <i>E. Coli</i></a>	Screen out <i>E. coli</i> Sequences
<a href="#">Screen Vector</a>	Screen out Vector Sequences
<b>Genome Resources</b>	
<a href="#">GMOD/GBrowse</a>	GBrowse local data genome browser
<a href="#">GRID Database</a>	Genomic Resource and Information Database Complete Tandem repeat analysis for human genome and other linked resources.
<a href="#">UCSC Mirror</a>	Mirror of popular UCSC genomic (human and mouse) genome browser
<b>General Tools</b>	
<a href="#">SeqWeb</a>	Web Interface to GCG (Local Access)
<a href="#">SWISS-MODEL</a>	SWISS-MODEL is an Automated Protein Modelling Server running at the ABCC.

### Bioinformatics/Proteomics Software

**Sequence Analysis** - Geospiza Finch LIMS System, Wisconsin Package (GCG), Clustal, Docsearch, Fasta, Blast, Mfold, Boxshade, Dotter, Genscan, Genehunter, Mview, Scan\_for\_Matches, Primer, Silent, Phred/Phrap, PolyPhred, Consed, RepeatMasker, Crossmatch, Staden, Readseq, Homochart, Sputnik, Profit, DSSP, Xtlsstr, MakePatterns, SeqDiff, FindStat, GelMap, EPCR.

**Linkage Analysis** - SAGE, Linkage, Simlink, Slink, FastLink, Vitesse, Sperm, Spermseg, Gasp, Rhmap, Sib-Pair.

**Phylogenetic Analysis** - Phylip, Paup, PAML, Puzzle, Adaptsite, Molphy, Phylowin, Lintre, Mismatch.

**Databases** - Genbank, GenPept, PIR, SwissProt, NRL\_3d, Prosite, PDB, NRDB-nucleic, NRDC-protein.

### How to Contact Us and Access Applications

For inquiries about computer services, please contact the ABCC at 301-846-5763,  
 FAX : 301-846-5762,  
 E-mail : [helpuser@ncifcrf.gov](mailto:helpuser@ncifcrf.gov)  
 URL : <http://www.abcc.ncifcrf.gov>