

NUS HPCBIO Portal

Grace Foo
(HPC, Computer Centre)

The HPCBIO portal offers life sciences researchers convenient access via the web to over 30 applications for Bioinformatics and Molecular Modelling. Each application has a graphical point and click interface for easy input of application and job parameters. The user does not need to know details such as hostnames and queues as the Portal knows how to direct submitted jobs to appropriate resources located at Computer Centre, BIC (Bioinformatics Center) and other centres at NUS.

The Portal is designed to provide an integrated environment for users to submit, and manage their jobs and their home / work space. It is menu driven and with the click of mouse, selections are made with minimal typing. Through these, we hope to take user experience to higher levels of productivity and convenience. In this article, we take a look at some Portal features in greater detail.

Who may use HPCBIO

All NUS students, staff and researchers who need large computational power for their research projects may use the portal after first registering for a HPC account with Computer Centre. Users login to the portal with their NUSNET user ID and password. On logging in, they will be able to see the home directory in their HPC account.

Applications available

Some 30 popular Life Science applications are accessible from the Portal. Most of the applications are open sourced, with a few commercial applications. Approximately a quarter of the applications are parallelised and run on HPC clusters. The following table lists the applications:

Category	Description	Applications/tools
Docking	Tools to predict how small molecules, such as drug candidates, bind to a receptor of known 3D structure.	Autodock
Modelling & structure	A range of techniques to model or mimic the behavior of molecules including atomistic level description of the molecular systems.	Amber, NAMD, CHARMM, Gromacs, Modeller, Gaussian
Phylogeny	Tools to study the evolutionary	FastDNAmI,

	relatedness among various groups of organisms.	Phase*, Phylip* *implemented in phase 2
Sequence alignment and analysis	Tools that arrange the primary sequences of DNA, RNA, or protein to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships between the sequences. For pair wise or multiple sequence alignment.	FastA, ClustalW-MPI, Muscle, Blast, Emboss, MGAlign, SMS, SRS, Meme/Mast DNAFSminer
Protein classification, analysis and modelling	Miscellaneous protein tools located at BIC.	CysView, NetCTL, SCOP, SDPMOD, SVMProt, VAGAT, HMMER
Miscellaneous	Miscellaneous formatting and graphing tools located at BIC.	ESript, SGM

The applications are categorised and may be found in the appropriate tabs as seen below. The input/output files, application and job parameters may be easily selected from the menus. The user need not specify a host or queue to run on but he may need to specify the application running mode – serial or parallel, if both are available.

The screenshot shows a web portal with a navigation menu at the top containing tabs for Home, File Manager, Job Manager, Docking, Modelling & Structure (highlighted), Phylogeny, and Sequence Alignment & Analysis. Below these are sub-tabs for Protein Classification and Miscellaneous. A secondary menu lists applications: AMBER, Gaussian03, Modeller, NAMD, CHARMM, and GROMACS. The main content area is titled 'Gaussian03 Portlet' and contains the following text: 'Gaussian 03 predicts the energies, molecular structures, and vibrational frequencies of molecular systems, along with numerous molecular properties derived from these basic computation types. It can be used to study molecules and reactions under a wide range of conditions, including both stable species and compounds which are difficult or impossible to observe experimentally such as short-lived intermediates and transition structures.'

Below the text is a 'Mandatory Inputs' form with the following fields:

- Choose G03 Type: Serial (dropdown menu)
- Location of Input File: (eg. /home/svu/t0700655/test100.com) [text input] [Browse button]
- Output filename: (eg. test100.out) [text input]
- Job Description (Optional) [text input]
- [Submit button]

On submission of the job, the Portal returns the user with a Job ID which enables him to check the job status in the Portal's Job Manager.

Home | File Manager | Job Manager | Docking | **Modelling & Structure** | Phylogeny | Sequence Alig
 Protein Classification | Misc

AMBER **Gaussian03** Modeller NAMD CHARMM GROMACS

Gaussian03 Portlet

Your Job Submission Result:

JOBID = Job Submission Succeeded, your JOBID: 432241

Your output files will be under /home/svu/t0700655/gaussian

The Job Manager communicates with the batch job scheduling system which returns the job status. When the status is DONE, the job has been completed successfully.



Home | File Manager | **Job Manager** | Docking | Modelling & Structure | Phylogeny | Sequence Alignment & Analysis
 Protein Classification | Miscellaneous

Job Manager

Job Manager Portlet

HPC Job Manager Web Interface: View/Manage your HPC Job Submissions

View, Manage Jobs

Date	JobID	Application	Job Description	Status
2009-09-17	436851	-	Submitted through command line	DONE
2009-09-17	436843	-	Submitted through command line	DONE
2009-09-16	436615	-	Submitted through command line	DONE
2009-09-16	436607	-	Submitted through command line	DONE
2009-09-16	436567	-	Submitted through command line	DONE
2009-09-16	436409	-	Submitted through command line	DONE
2009-09-16	436396	-	Submitted through command line	EXIT
2009-09-16	436355	-	Submitted through command line	DONE
2009-09-15	435753	ClustalW-mpi	ClustalW-mpi Job	DONE
2009-09-10	432241	GAUSSIAN03	Gaussian03 Job	DONE
2009-08-25	417370	-	Submitted through command line	DONE
2009-08-25	416793	AMBER	AMBER Job	DONE

Files, work space and the File Manager

Through the Portal's File Manager, the user may check the results / output of the completed job. To navigate his directory and work with files, basic functions are available such as file editing, creation, deletion, copy and directory listing, creation, deletion, copy. There are also some special functions such as file zip, download and upload for users to easily transfer files to/from their local machine

to their HPC space. Users may access their HPC home directory or the temporary directory.

Home **File Manager** Job Manager Docking Modelling & Structure Phylogeny Sequence Alignment & Analysis
Protein Classification Miscellaneous

File Manager




FileManager Portlet

Browse... Upload File to Server (File size max. 100MB)

Choose Base

Directory : /home/svu/t0700655

Current Working Directory: /home/svu/t0700655

Select	Type	Filename	Size(bytes)	UID	GID	Date Modified
<input type="checkbox"/>		Amber	1024	11084	300	Tue Aug 25 15:22:50 SGT 2009
<input type="checkbox"/>		amber	2048	11084	300	Thu Mar 19 17:42:49 SGT 2009
<input type="checkbox"/>		Amber10	1024	11084	300	Tue Aug 25 15:22:49 SGT 2009

Weblinks

Portal login <https://hpcbio.nus.edu.sg/gridsphere/gridsphere>

User Guide: <https://inetapps.nus.edu.sg/comcen/hpcbio/portal/userguide.html>

FAQ: <https://inetapps.nus.edu.sg/comcen/hpcbio/portal/faq.html>

The Portal was launched on 15th October. Do look out for seminars and application training that we will announce in the coming year.

For questions or feedback, please contact the author at ccefoog@nus.edu.sg