Vigyaan: An Integrated Software Workbench for Bioinformatics and Computational Biology

Pratul K. Agarwal,1,2 David Jung,2 Al Geist1,2
1Computational Biology Institute; 2Computer Science and Mathematics Division, Oak Ridge National Laboratory, Bethel Valley Road, Oak Ridge, TN USA

Project web-site: http://www.vigyaanCD.org/ E-mail contact: agarwalpk@ornl.gov

Computational analysis of genomic data and molecular modeling of biomolecular systems are becoming an integral part of an increasing number of biological investigations. An enormous amount of data is currently being generated and routinely analyzed. Software tools required for analysis and modeling of biological data also continue to grow at a fast pace. Researchers are spending a considerable amount of time and resources in locating, obtaining and installing these software tools. Moreover, time spent in learning new software is also becoming non-trivial as software packages continue to grow more complex.

Bioinformatics and biomolecular modeling software tools have different formats, requirements and availability, as they are developed by a diverse community of researchers and programmers. In recent years, PCs have become ubiquitous and, therefore, have emerged as a popular low cost hardware choice. On the software front GNU/Linux has emerged as a low cost and high performance option.1,2 However, scientific software has not kept up with PC hardware and software development. Available bioinformatics and biomolecular modeling software requires a diverse set of supplementary software libraries, making it a challenging, and sometimes frustrating, task for the end user. Further, the user is unable to take full advantage of applications due to lack of interoperation between different software packages.

Vigyaan is an integrated software workbench for bioinformatics and computational biology. Bootable VigyaanCD can be prepared easily from the ISO images available from the project homepage. VigyaanCD consists of a complete computer operating system based on the Debian GNU/Linux system and boots a PC into the workbench environment. The booting time on current PC hardware ranges between 2 to 5 minutes, thus providing a workbench with ready-to-use software in extremely short duration without having to install software or alter data on the local hard disk of the PC. The applications are easily accessible through command line and graphical interface (where available).

Currently, over 25 applications are included on VigyaanCD. These applications range from genomic analysis, biomolecular structural visualization and manipulation to molecular modeling of biomolecular systems. In addition, tools for data analysis visualization, numerical and statistical calculations, image creation and editing are also available. Table 1 provides a list of scientific software available on VigyaanCD; further details are available on the project web-site. VigyaanCD also includes genomes from *Escherichia coli* (K 12) and *Synechococcus* sp. as samples. An additional feature of Vigyaan is that it provides C/C++/Fortran compilers, Java runtime environment, and tools required to compile and install other software. Other applications for which source code is available can be easily compiled and installed on any PC, independent of its host operating system. New applications can also be developed and tested in this environment. Storage space for Vigyaan workbench is limited by default as the workspace is created in the memory and therefore lost after rebooting the computer. However, it is possible to use the local hard disk to create extended storage. PC clusters based on commodity hardware are increasingly being used to perform bioinformatics and computational biology calculations. Vigyaan is also suitable for booting a collection of PCs to perform as a cluster, enabling high-throughput analysis of genomic and sequence data and detailed molecular modeling.

We are also developing a more advanced workbench called BiLab. BiLab is a completely integrated workbench for bioinformatics and computational biology and will work on Windows and GNU/Linux platforms. Enhancing the advantages of Vigyaan, BiLab will offer an integrated environment where additional software can easily be installed; workspace and user data can be permanently stored on the local disk.

### Availability
VigyaanCD images are being distributed free of cost under GNU’s General Public License (GPL).5 They are available from the project web-site: http://www.vigyaanCD.org/ and many mirror web-sites around the world. Most of the software packages included on VigyaanCD are also licensed under GPL. However, additional software packages are available through special license agreement.

<table>
<thead>
<tr>
<th>Bioinformatics</th>
<th>Artemis, Bioperl, ClustalX, Cn3D, EMBOSS, NCBI Toolkit, NJPlot, SMILE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computational biology (molecular modeling)</td>
<td>Garlic, GROMOS, Open Babel, PyMOL, Rasmol, Raster3D, TINKER</td>
</tr>
<tr>
<td>Computational chemistry</td>
<td>Ghemical, J Mol, MPQC, PSI3, XDrawChem</td>
</tr>
<tr>
<td>Other tools</td>
<td>GNU C/C++/Fortran compilers, GNU R, Gnuplot, Octave</td>
</tr>
</tbody>
</table>