

Open discovery: An integrated live Linux platform of Bioinformatics tools

Umashankar Vetrivel¹ * and Kalabharath Pilla¹

¹Department of Bioinformatics, School of Biosciences, SRM University, Ramapuram, Chennai, India; Umashankar Vetrivel* - Email: vumashankar@gmail.com; * Corresponding author

received October 15, 2008; accepted October 30, 2008; published December 06, 2008

Abstract:

Historically, live linux distributions for Bioinformatics have paved way for portability of Bioinformatics workbench in a platform independent manner. Moreover, most of the existing live Linux distributions limit their usage to sequence analysis and basic molecular visualization programs and are devoid of data persistence. Hence, open discovery – a live linux distribution has been developed with the capability to perform complex tasks like molecular modeling, docking and molecular dynamics in a swift manner. Furthermore, it is also equipped with complete sequence analysis environment and is capable of running windows executable programs in Linux environment. Open discovery portrays the advanced customizable configuration of fedora, with data persistency accessible via USB drive or DVD.

Availability: The Open Discovery is distributed free under Academic Free License (AFL) and can be downloaded from <http://www.OpenDiscovery.org.in>.

Keywords: Open discovery; Bioinformatics Linux distribution; Bioinformatics tools; Bioinformatics software; molecular modeling; molecular dynamics; docking; sequence analysis; structural Bioinformatics

Background:

The Linux computing technology now enables users to boot Linux from USB flash drives and CDs/DVDs in a form called as LIVE media, apart from access via standard installation methods. Such technology gives mobility, portability, personalization and ease of usage. Linux has been highly sought of for its versatility to customize and availability as open source.

Open Discovery is one such customized Linux distribution developed based on RedHat's Fedora [1] and is capable of booting from USB flash drives, Live-DVDs and Hard-disks, superseding traditional knoppix formats [2] in terms of high speed performance and customization. It comprises of a comprehensive collection of bioinformatics software packages ranging from basic sequence analysis to computationally complex processes like molecular dynamics, molecular modeling and Docking. Open Discovery is best suited for current generation of multicore computer systems.

Methodology:

Open Discovery is a Fedora based Linux distribution, built on latest release "Fedora sulphur 9" [3]. Creation of Open

Discovery is accomplished by shell scripting, customization of software packages, creating binaries of software packages and usage of Livecd-tools [4] on a Linux platform.

Features and functionalities

The salient feature of Open Discovery is its ability to boot from a USB flash drive with data persistence. This enables the user to carry the operating system in an USB drive along with the data generated during analysis, and is non-destructive on reboot. This feature is enabled by retaining the snapshot of the live data image on the USB flash drive itself, rather than in RAM as in case of traditional knoppix formats.

Open Discovery is created to form a customizable and portable bioinformatics platform, integrating a comprehensive range of bioinformatics software. The platform features genuine opensource programs like, EMBOSS [5] for sequence analysis, Jackal [6] for Protein Structure prediction, patchdock [7] for Protein-Ligand Docking and Gromacs [8] for Molecular Dynamics. The complete list of tools available in Open Discovery can be accessed from the homepage [9]. A snapshot of few applications in Open Discovery is shown in figure 1.

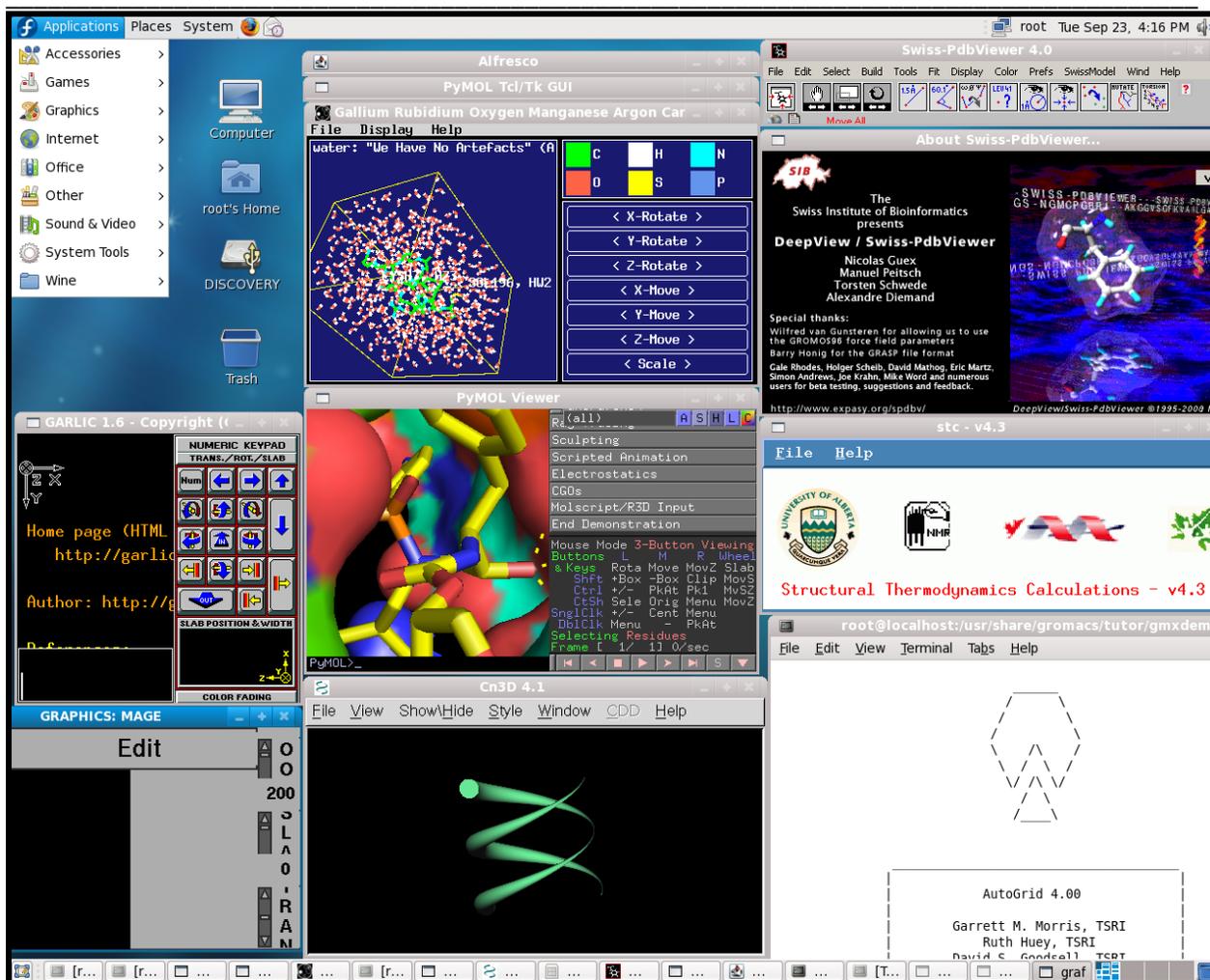


Figure 1: Open discovery desktop showing few Bioinformatics applications.

Apart from bioinformatics applications, Open Discovery also comprises of many other features like, writefst program enabling the users to install software while the liveDVD is in active mode, enforced SELinux with security features, compilation of best of breed software on the media, easy to craft livedvd with customized repositories, packages and art work and supports a wide variety of hardware.

Open Discovery also encompasses Wine [10], a translation layer capable of running windows applications on Linux and other POSIX compatible operating systems. This feature allows the user to install and run bioinformatics tools which are designed to run on Microsoft's Windows operating systems on Open Discovery environment.

The usage of Open Discovery as a LiveUSB with persistent data overlay enables the users to create new login accounts and passwords, to install additional software and to update the operating system. Hence, overcoming the restricted

software installation and usage as featured in distributions like BioKnoppix, DNA Linux etc.

Apart from the usage as LiveUSB, Open Discovery can also be used as a Live-DVD. The user can easily access the NTFS/FAT formatted windows drives from the Linux environment on a single mouse-click.

Caveats and future development:

The Graphical user interface may not be able to function of certain laptops and desktops which use onboard GPU in nvidia chipset motherboards. However, there are no issues on systems which have dedicated nvidia graphics cards. The current version doesn't boot on PowerPcs (Macs) and on systems other than Intel and AMD.

Currently 32-bit version is released and a 64 bit version is under development. Future version of Open Discovery is being developed to facilitate cluster computing with Graphical user

interface tools and parallel implementation of Docking and Molecular Dynamics tools.

References:

- [01] <http://fedoraproject.org>
[02] <http://bioknoppix.hpcf.upr.edu>
[03] <http://fedoraproject.org/wiki/Release/9/Alpha/ReleaseNotes>
[04] <https://admin.fedoraproject.org/pkgdb/packages/name/liquidcd-tools>
[05] <http://emboss.sourceforge.net/>
[06] http://wiki.c2b2.columbia.edu/honiglab_public/index.php/Software:Jackal_General_Description
[07] <http://bioinfo3d.cs.tau.ac.il/PatchDock>
[08] <http://www.gromacs.org>
[09] <http://OpenDiscovery.org.in/Documentation.html>
[10] <http://www.winehq.org>

Edited by P. Kanguane

Citation: Umashankar & Kalabharath, Bioinformatics 3(4): 144-146 (2008)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.