

Installation Instructions for *xcms*

Colin A. Smith

October 1, 2012

Introduction

This document describes how to install *xcms* and, if necessary, also obtain and install R. The *xcms* package includes C code which needs to be compiled, and also uses the NetCDF library for reading AIA format NetCDF mass spectral data files. Pre-compiled binaries which include the NetCDF library are available for Windows and Mac OS X. Users of Linux and other platforms must install the NetCDF library themselves.

The *xcms* package and a demonstration data package, *faahKO*, are currently available from several sources. The first is the web site of the Bioconductor open source software project. (<http://www.bioconductor.org/>) The second is the web site of the METLIN Metabolite Database. (<http://metlin.scripps.edu/download/>) Both sites contain source and binary distributions, although due to differences in distribution, one may host a slightly more up-to-date version than the other.

R is available for download through the Comprehensive R Archive Network (CRAN). Visitors are encouraged to use one of the many local mirrors of the CRAN site for efficient downloading. (<http://cran.r-project.org/mirrors.html>) While the vignettes included with *xcms* give many examples of R use and syntax, new users are strongly encouraged to skim some of the introductory material in the Manuals section of the R web site. (<http://www.r-project.org/>) “An Introduction to R” is an especially useful starting point.

1 Windows Installation

To make installation as straightforward as possible for all users, *xcms* includes a binary version of the NetCDF library in the `inst/netcdfdll` subdirectory. The version currently included is 3.6.1-beta1, as available from the Unidata web site¹.

1. Download and install the current version of R. Detailed instructions for doing so are available on the CRAN site.

¹<http://www.unidata.ucar.edu/packages/netcdf/>

2. Once you have R installed, launch it and select Bioconductor repository using the `Packages > Select repositories...` menu item.
3. Use the package installer to automatically download and install *multtest* with the `Packages > Install package(s)...` menu item.
4. Download the *xcms* and *faahKO* Windows binaries from one of the sources listed in the introduction.
5. Install *xcms* and *faahKO* using the `Packages > Install package(s) from local zip files...` menu item.

2 Mac OS X Installation

The Mac OS X binaries of *xcms* may be installed without the Developer Tools and without a separate installation of the NetCDF library. Its code is pre-compiled and statically linked against the NetCDF library version 3.6.0-p1. If you wish to compile it yourself, you may obtain the library from software distribution projects such as Fink² or DarwinPorts³. Alternatively, you may compile it yourself using the instructions provided below.

1. Several of the functions in *xcms* which generate PNG images require an X11 display device. If it is not already installed, you will need to install X11⁴.
2. Download and install the current version of R. Detailed instructions for doing so are available on the CRAN site.
3. Download the *xcms* and *faahKO* Mac OS X binaries from one of the sources listed in the introduction.
4. Launch R and open the R Package Installer using the `Packages & Data > Package Installer` menu item. Select the “At User Level” radio button.
5. Select “BioConductor (binaries)” from the pop up menu and use the package installer to automatically download and install *Biobase* and *multtest*.

If Bioconductor is preparing for its next release and binaries aren’t yet available for your version of R, you will need to install the Apple Developer Tools⁵ to allow installation of source Bioconductor packages. Use the “BioConductor (sources)” option instead.

²<http://fink.sourceforge.net/>

³<http://netcdf.darwinports.com/>

⁴<http://www.apple.com/downloads/macosx/apple/x11formacosx.html>

⁵<http://connect.apple.com/>

6. Select “Local Source Package” from the popup menu and install the *xcms* and *faahKO* packages you previously downloaded.

3 Obtaining the NetCDF Library

If you are using Linux or another UNIX-like operating system, you will have to obtain the NetCDF library before installing and using *xcms*. Many Linux distributions include that package so in most cases, locating and installing the version that comes with your distribution will be the best option. However, you may also install the NetCDF package using the following instructions:

1. Download the gzipped tar file of the NetCDF source (`netcdf.tar.gz`) from the Unidata web-site¹. Then extract the archive and change to the source directory.

```
tar -xzf netcdf.tar.gz
cd netcdf-*/src
```

2. Configure, compile, and install the library. By default it is installed in the directory you extracted in the previous step. If you wish to install it in another location, use the `--prefix` argument with the configuration script.

```
./configure --prefix=/usr/local
make
make install
```

It is important to note that, when compiled for the x86_64 architecture, the NetCDF library must be compiled with the `-fPIC` flag to produce position-independent code. Certain Linux distributions, such as SuSE Professional 9.2, do not yet enable that option when packaging the NetCDF library. In that case you must compile and install the library yourself using the instructions above. To enable that flag, set the `CFLAGS` environment variable prior to compilation.

```
export CFLAGS=-fPIC (for sh, bash, etc.)
setenv CFLAGS -fPIC (for csh, tcsh, etc.)
```

4 Obtaining the *rgl* Package

Though not listed in any of the dependencies of *xcms*, the *rgl* package does get limited use. It provides an interface for creating interactive, 3D graphics using OpenGL. The *rgl* package is currently under development and does not yet provide sufficient functionality or stability to warrant its listing as an official dependency. It is used by a single

method, *plotSurf*, for plotting out a 3D surface representation of part of an LC/MS or GC/MS experiment. Due to limitations in *rgl*, axis labels are currently not implemented. However, it can be instructive for visualizing a small number of peaks in three dimensions.

The best place to obtain source and Windows binary packages of *rgl* is directly from CRAN. Mac OS X binaries are available in the same place you download the R installer.