Set-up for

BRC hands-on workshops

Preview video tutorials
Configure your laptop
Download updates
Know your computer OS
Register for IGV

Training materials for Next Gen Sequencing Analysis @ BRC

Jean-Yves Sgro
With Jeremy Glasner & Xiao-yu Liu

Version 3.2.0
Image credits:

Tux: the Linux mascot: https://en.wikipedia.org/wiki/Tux
146 base pairs DNA portion of a nucleosome structure (PDB 1aoi) rendered with PyMOL

Workbook – © 2016 – All Rights Reserved

Corresponding author: Jean-Yves Sgro - jsgro@wisc.edu

Instructor(s):

Jean-Yves Sgro, Ph.D.

Associate Scientist, Biotechnology Center and Genomics Center of Wisconsin.

Office: 4434 Genetics-Biotech Center

Email: jglasner@wisc.edu

Class location:

Genetics-Biotechnology Center Building - 425 Henry Mall

Campus Map short link:

http://bit.ly/1m9n5OZ

Google map short link:

https://goo.gl/maps/BDNQ0
This Book Belongs to:

Name: __________________________

Cluster User Name: _______________

Note: for safety reasons student user names will be disabled after class.

Cluster Password: _______________

Note: for safety reasons password will be reset after class.

Assigned Cluster For the Exercises: __________________________

(e.g. agar1, agar2, agar3, sumo)
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreword &amp; Summary</td>
<td>5</td>
</tr>
<tr>
<td>Bring a Laptop</td>
<td>6</td>
</tr>
<tr>
<td>1. Laptop required</td>
<td>6</td>
</tr>
<tr>
<td>1.1. Bring your own</td>
<td>7</td>
</tr>
<tr>
<td>1.2. Free Library Loan</td>
<td>7</td>
</tr>
<tr>
<td>1.3. Rental</td>
<td>8</td>
</tr>
<tr>
<td>1.4. Wireless</td>
<td>9</td>
</tr>
<tr>
<td>1.5. Bring your power supply!</td>
<td>9</td>
</tr>
<tr>
<td>2. Admin password</td>
<td>9</td>
</tr>
<tr>
<td>Software required for all computers</td>
<td>10</td>
</tr>
<tr>
<td>Software required if you have Mac</td>
<td>11</td>
</tr>
<tr>
<td>Quartz</td>
<td>12</td>
</tr>
<tr>
<td>Software required if you have a PC</td>
<td>13</td>
</tr>
<tr>
<td>Software required for a Linux system</td>
<td>15</td>
</tr>
<tr>
<td>Unix Shell</td>
<td>16</td>
</tr>
<tr>
<td>1. Required study: the Unix Shell</td>
<td>17</td>
</tr>
<tr>
<td>2. Optional</td>
<td>19</td>
</tr>
<tr>
<td>Register on IGV genome viewer web site</td>
<td>21</td>
</tr>
<tr>
<td>Know Your Computer OS</td>
<td>23</td>
</tr>
<tr>
<td>1. Operating System of your computer</td>
<td>23</td>
</tr>
<tr>
<td>2. File name extension: the “real name” of files</td>
<td>23</td>
</tr>
<tr>
<td>3. Macintosh</td>
<td>24</td>
</tr>
<tr>
<td>4. Windows</td>
<td>25</td>
</tr>
<tr>
<td>4.1. For Windows Vista, Windows 7, and Windows Server 2008</td>
<td>25</td>
</tr>
<tr>
<td>4.2. For Windows 2000, Windows XP, and Windows Server 2003</td>
<td>26</td>
</tr>
<tr>
<td>4.3. Control Panel (Windows 10)</td>
<td>26</td>
</tr>
</tbody>
</table>
This booklet contains information to prepare for attending a BRC workshop at the Biotechnology Center

**SUMMARY**

1) **Bring a computer notebook**, Mac or PC
2) **Update Java:**
   
   https://java.com/en/
3) **PC**: install MobaXterm
   
   http://mobaxterm.mobatek.net/download.html
4) **Mac**: install X11
   
   http://xquartz.macosforge.org/
5) **Register to use IGV**, the Integrated Genome Viewer
   
   http://www.broadinstitute.org/igv/
6) **View videos about "Unix Shell"**
   
Bring a Laptop

Summary:
Bring a wireless capable laptop & power supply

Workshops are geared to using BRC Linux computers for Next Gen Sequencing analysis. As such, any local computer can be used as a connecting terminal to link to the BRC nodes.

New class format is now assuming that you bring your own laptop to class. Follow these instructions to be ready for class.

1. Laptop required

Classes are no longer taught in dedicated computer classrooms and you are required to bring a computer with you.
1.1. Bring your own

Your own computer can be a Mac, a PC or a Linux-based computer. The computer will only be used as a terminal to access the remote BRC Linux system and run simple visualization software.

1.2. Free Library Loan

UW Libraries have computers to loan, up to 20 times per semester without cost as far as we can tell.

Locations and numbers of available computers are visible on a map at this link:

https://ecs.library.wisc.edu/

The Steenbock Library is one possible loan location that might be closer to most life scientists:

http://www.library.wisc.edu/steenbock/services-at-steenbock/computers/laptops/

or short URL: http://bit.ly/1jdLcw2

Laptops are available for checkout from the Circulation Desk near the entrance on 2nd floor.

Call 608-262-1371 for more information.

Laptops
50 Macbook Air and 5 Macbook Pro laptops

- Check availability
- Runs both Windows and Apple OS.
- Includes carrying case and power adapter. Loaded with standard software including Microsoft office. Note: Adobe Creative suite is only available on Infolab Desktop computers.
Laptop checkout policies

- Steenbock Infolab laptops may not be reserved in advance.
- Laptops are checked out at the Circulation/Reserves desk.
- Two Forms of ID required:
  - UW-Madison ID
  - Driver’s license, Wisconsin ID, or passport
- Loan period is 3 days

1.3. Rental

It is possible to rent a computer for $10 for the day at DoIT. See [https://it.wisc.edu/services/rentals/](https://it.wisc.edu/services/rentals/)

The information from that page is reproduced here:
(Last verified 2/2016.)

Notebook computers for rent

- **Apple MacBookPro (dual boot):** 2.26GHz, 2-4GB RAM, Mac OS X 10.9 with Microsoft Office 2010, Windows 7 with Microsoft Office 2013, Google Chrome, Mozilla Firefox, Skype and more.

- **Dell Latitude E6400:** 2.66GHz, 4GB RAM, Windows 7 with Microsoft Office 2013, Google Chrome, Mozilla Firefox and more.

- **Software & data policy:** Customers are free to install any other software that they require. Please note that we do not support customer-installed software, and all software and data will be erased from the computer at the end of the rental period.

<table>
<thead>
<tr>
<th>Model</th>
<th>Daily</th>
<th>Weekly (7 Days)</th>
<th>Monthly (4 weeks)</th>
<th>Semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notebook computer</td>
<td>$10</td>
<td>$25</td>
<td>$75</td>
<td>$150</td>
</tr>
</tbody>
</table>
**Rental inquiries:**

For availability information, stop by our location or email us at rentals@doit.wisc.edu. Email inquires and rental reservations will receive an email response within four business hours.

### 1.4. Wireless

![WiFi symbol](image1)

Although it is standard now, the computer must have standard wireless capabilities.

### 1.5. Bring your power supply!

Power-strips may be available or plug directly in the wall.

### 2. Admin password

You should have Administrative privileges to install software on the computer you will bring in class.
Software required for all computers

Summary:
Install or update Java.

✔ TASK: go to java.com and install/update Java for your computer.

Java needs to be installed or updated as it is required to run the IGV genome viewer.
Software required if you have Mac

Summary:
Install or update X11/XQuartz

Modern Macintosh computers running Mac OS X are already “Unix-based” computers and most software is already present.

Depending on the version of Mac OS the following software may be useful, but probably dispensable: **X11**

From the Apple web site:

X11 is no longer included with OS X, but X11 server and client libraries for OS X are available from the XQuartz project: [http://xquartz.macosforge.org](http://xquartz.macosforge.org). You should use the latest available version of XQuartz.

On the Quartz web (reproduced below) download the .dmg file and
A version of the X.Org X Window System that runs on OS X

The XQuartz project is an open-source effort to develop a version of the X.Org X Window System that runs on OS X. Together with supporting libraries and applications, it forms the X11.app that Apple shipped with OS X versions 10.5 through 10.7.

Quick Download

<table>
<thead>
<tr>
<th>Download</th>
<th>Version</th>
<th>Released</th>
<th>Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>XQuartz-2.7.8.dmg</td>
<td>2.7.8</td>
<td>2015-10-17</td>
<td>For OS X 10.6.3 or later (including El Capitan)</td>
</tr>
</tbody>
</table>

A list of all available XQuartz releases can be found here http://xquartz.macosforge.org/trac/wiki/Releases

Note: XQuartz will be installed in /Applications/Utilities on your computer. You may need Admin privileges to accomplish the installation.

Note: When launching XQuartz a terminal of type “xterm” might appear. In general this is not the preferred terminal as it is difficult to copy/paste text commands. The terminal used in class is: /Applications/Utilities/Terminal and can easily copy/paste plain text within.
Software required if you have a PC

Summary:
Install MobaXterm

If you have a Windows-PC it should be running either Windows 7, 8.1, or later as Windows XP is no longer allowed on UW networks.

The following, free software contains all that is need to connect and display data from a remote Linux computer and should be installed before class: MobaXterm

✔ TASK  Go to mobaxterm.mobatek.net

Click on Download and Download and install the “Free download” Home Edition.
The features labeled “unlimited” in the paid edition do not play a role in how we use the software and therefore **the free edition is perfectly fine**.

Click Download now:

There are 2 installer versions:

**Portable edition**

**Installer edition**

The final software is the same. The difference is a “Windows thing” and resides only on the installation method. The portable version, once no longer needed can be simply placed into the recycle bin. The installed version would have to be removed via the Windows-specific method to “Add/Remove” software via the control panel.

✔️ **TASK** Install the software and explore the menu options.
Software required for a Linux system

A Linux laptop should already have everything necessary and nothing further should be required.

Make sure that you know where to find the “Terminal” or “Console” for text commands.

Verify that your Java is up to date (see above) and that you know the administrative password.
Unix Shell

Summary:
Go to software-carpentry.org/lessons.html to learn about The Unix Shell online before class.

This is the most important preparation!

Required before attending class!

Next Gen Sequencing analysis at BRC is performed on Linux computer systems. Therefore understanding basic Linux commands at the command-line is a requirement for all workshops. It is similar to the idea of knowing the alphabet before learning to write.

To save time in class with the minute details of commands, it will now be expected that students will review an online tutorial on the subject. The material will be reviewed in class in a fraction of the time allotted before, therefore providing more time for actual exercises since the basic learning will have occurred beforehand.
Specific resources will be provided below with additional, optional resources for UW.

1. Required study: the Unix Shell

This subject is an essential part of using any Linux/Unix system and line-command software. The following resource is free and accessible by all:

![Software Carpentry Logo](software-carpentry.org)

**software-carpentry.org**

**TASK** Go to software-carpentry.org/lessons.html

There are 2 versions available:

**Video format (Version 4) (recommended)**
permanent archive: http://bit.ly/1Mf5Od0

**Text format (Version 5.3)**
http://swcarpentry.github.io/shell-novice/

The content is similar but not identical. Choose the learning method you prefer. However:

The **video version** (version 4) is probably easiest to follow and is therefore **recommended**.

The chapters for the video version are as follows (including here the direct YouTube links:)

---
The last 2 entries here have been inverted compared to the Version 4 web page as it seemed more appropriate in this order. They are listed here as well as the direct YouTube link in case the version 4 web page becomes unavailable.

<table>
<thead>
<tr>
<th>“The Unix Shell”</th>
<th>YouTube Link</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td><a href="http://youtu.be/U3iNcBtycaQ">http://youtu.be/U3iNcBtycaQ</a></td>
<td>4:10</td>
</tr>
<tr>
<td>Files and Directories</td>
<td><a href="http://youtu.be/qrkvcX4HmYo">http://youtu.be/qrkvcX4HmYo</a></td>
<td>9:56</td>
</tr>
<tr>
<td>Creating and Deleting</td>
<td><a href="http://youtu.be/FmXtSATyuII">http://youtu.be/FmXtSATyuII</a></td>
<td>6:26</td>
</tr>
<tr>
<td>Pipes and Filters</td>
<td><a href="http://youtu.be/5crQ2hMCgcs">http://youtu.be/5crQ2hMCgcs</a></td>
<td>9:12</td>
</tr>
<tr>
<td>Permissions</td>
<td><a href="http://youtu.be/yFMLnu5HWgM">http://youtu.be/yFMLnu5HWgM</a></td>
<td>10:55</td>
</tr>
<tr>
<td>Finding Things</td>
<td><a href="http://youtu.be/7uZW1Fh8JLY">http://youtu.be/7uZW1Fh8JLY</a></td>
<td>9:23</td>
</tr>
<tr>
<td>Job Control</td>
<td><a href="http://youtu.be/3XJHXLr7XA">http://youtu.be/3XJHXLr7XA</a></td>
<td>5:38</td>
</tr>
<tr>
<td>Variables</td>
<td><a href="http://youtu.be/faAvqcWD11c">http://youtu.be/faAvqcWD11c</a></td>
<td>6:50</td>
</tr>
</tbody>
</table>

**TOTAL:** 1h25:19

It should take less than 2 hours to go over all the videos.
The material presented in these videos corresponds to basic knowledge that will be USED IN CLASS.

Your ability to follow in class will be impacted by what you learn ahead of time with this material.
Students that have not followed these videos will find it hard to follow the in-class review. Therefore, make all best effort to see these videos before class.

2. Optional

Video classes are also offered on Lynda.com (available to all UW personnel and requires NetID login.)

Using the “Search” field to search for keywords can help find relevant classes. Below are a few examples pertinent to this workshop:

Search for the word “bash” – currently yields one class:

<table>
<thead>
<tr>
<th>Course Title</th>
<th>Description</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up and Running with Bash Scripting</td>
<td>A fast-paced introduction to Bash script, a popular shell scripting language on Mac OS X and Linux.</td>
<td>1h25min</td>
</tr>
</tbody>
</table>

This is perhaps the most relevant class on the Lynda.com site. However below are other useful entries:

Search first for the word “Unix” and a single lesson will appear:

<table>
<thead>
<tr>
<th>Course Title</th>
<th>Description</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unix for Mac OS X Users</td>
<td>Unlock the powerful capabilities of Unix that underlie Mac OS X, teaching how to use command-line syntax to perform common tasks such as file management, data entry, and text manipulation.</td>
<td>6h35min</td>
</tr>
</tbody>
</table>

You can also search for the word “Linux” and the search can be reduced further from clickable options on the left hand side of the results page (e.g. click on “Beginner.”)
The following entries could be useful for the Linux search. However, ALL of these are already BEYOND what is needed for this workshop!

<table>
<thead>
<tr>
<th>Course Title</th>
<th>Description</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up and Running with Ubuntu Desktop Linux</td>
<td>Learn how to install and set up Ubuntu as your main operating system or on a virtual machine, navigate the desktop, and work with popular Linux desktop programs.</td>
<td>1h42min</td>
</tr>
<tr>
<td>Up and Running with CentOS Linux with Scott Simpson</td>
<td>Learn how to install CentOS, perform common admin tasks from the command line, and get popular services running.</td>
<td>1h32min</td>
</tr>
<tr>
<td>AWK Essential Training with David D. Levine</td>
<td>Learn how to manipulate and format data with AWK, a Mac and Linux command-line tool.</td>
<td>2h01min</td>
</tr>
</tbody>
</table>

Notes on Lynda.com

STS [Software Training for Students] helps provide a campus-wide subscription to Lynda.com, giving you free access to a wide variety of training videos and class files. Trust us. It's awesome. Students, faculty, and staff can all use it.

All you need is your UW NetID to login and get access to over a thousand self-paced online courses.

Get more info at the following DoIT web page:

www.doit.wisc.edu/training/lynda/

Also available under “Services” within my.wisc.edu
Register on IGV genome viewer web site

The Integrative Genome Viewer (IGV) from the Broad Institute will be used to examine genome files created in class. Running a local version on your laptop rather than on the remote Linux computer will make the process faster and less prone to network problems.

Therefore, before class you please register on the web site (it’s free) to have access to the IGV software.

✔ TASK Go to www.broadinstitute.org/igv/
- Click Downloads (on the left side)
- Click on “Click here” to register
- Fill-in the requested info

Once registered, you only need your email (as your username) to access the software again from anywhere
Know Your Computer OS

In this section you will learn how to display the **real name** of a file in the computer that you are familiar with, Mac & PC. This may be important to transfer files to the BRC computers.

1. **Operating System of your computer**

Most likely you will access your BRC account from either a Windows-PC or a Macintosh computer. If you are using Linux as your main computer you might skip the Linux tutorial section.

Below are a few hints that may help in the process of accessing your account and transferring files with your own computer.

2. **File name extension: the “real name” of files**

✔ **READ:** By default, most modern computer will show file names without the traditional file name extension.
For example a document created with Microsoft Word may appear as “My Letter” in the directory where it is stored whether you are looking at it as an icon or a list. Most likely the real name of the file is either “My Letter.doc” or “My Letter.docx” depending on the version of Word that created it. In the same way, PDF documents will not show the “.pdf” and text files will not show the “.txt” filename extension etc. This was invented to “help” users but in some cases it is important or even imperative to know the exact name of the file to use.

File name before and after changes are made to show file extension

<table>
<thead>
<tr>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>My Letter</td>
<td>My Letter.doc</td>
</tr>
<tr>
<td>My Letter</td>
<td>My Letter.docx</td>
</tr>
<tr>
<td>My Letter</td>
<td>My Letter.pdf</td>
</tr>
<tr>
<td>My Letter</td>
<td>My Letter.txt</td>
</tr>
</tbody>
</table>

Before the extension is made visible, the only visual cue might be the icon shown for each file.

Here is how to show the real, complete name of your files on your computer:

3. Macintosh

- **General method:**
- Click anywhere on the Desktop to see the finder menu at the top of the screen
- On the upper left, next to the apple, click on “Finder” and choose “Preferences…”
- Click on the “Advanced” tab (top right of window)
- Check the box next to “Show all filename extensions”
- From now on all the file extensions will be shown.
• Single file method:

It is also possible to change this on a per file basis with the Finder menu Get Info (⌘ I) and clicking the square button Hide extension as shown in this example

4. Windows

Here is the information from Microsoft on this subject, explained here in term of Windows Explorer, but should work system wide:

From http://support.microsoft.com/kb/865219

Title of page: How to show or hide file name extensions in Windows Explorer

4.1. For Windows Vista, Windows 7, and Windows Server 2008

1. Start Windows Explorer, you can do this by opening up any folder.
2. Click Organize.
3. Click Folder and search options.
4. Click the View tab.
5. Scroll down until you notice Hide extensions for known file types, un-check this line by clicking the check box.
   Note To hide file name extensions, check this line.
6. Click OK
4.2. For Windows 2000, Windows XP, and Windows Server 2003

1. Start Windows Explorer, you can do this by opening up any folder.
2. Click Tools, and then click Folder Options.
3. Scroll down and then click Folder and search options.
4. Click the View tab.
5. Scroll down until you notice **Hide extensions for known file types**, un-check this line by clicking the check box.
   **Note** To hide file name extensions, check this line.
6. Click OK

4.3. Control Panel (Windows 10)

Some Windows versions might best be served by using the Control Panel, as is detailed here for Windows Vista and Windows 7 (switch with clickable arrow on right hand side of page):

[ or short URL: http://bit.ly/1MFBoVx ]

To show or hide file name extensions:

1. Open Folder Options by clicking the Start button 🌐, clicking Control Panel, clicking Appearance and Personalization, and then clicking Folder Options.

2. Click the View tab, and then, under **Advanced settings**, do one of the following:
   - To show file name extensions, clear the **Hide extensions for known file types** check box, and then click OK.
   - To hide file name extensions, select the **Hide extensions for known file types** check box, and then click OK.