Xpression Virtual System User Guide
Harwood Lab UW
March 2012

## Introduction

The Xpression graphical user interface is designed to be easy to use to quickly start analysing RNA-seq data. The goal of Xpression is to put the power of bioinformatic analysis into the hands of investigators with broad technical backgrounds. However, due to the nature of the supporting software, its requirement for a properly configured Unix-like operating system, and the complications that can arise, we have provided a variety of alternatives for getting Xpression onto your computer.

If you already have a Linux or Unix-like OS, your best option is to use the provided automated script that will install all required software from source. If your operating system is Windows or MacOSX we have provided a fully operational, system-independant graphical environment (Xpression VE) that can run Xpression without any further configuration.

The focus of this guide is the point-and-click installation of this environment onto your computer. It provides all the functionality to run Xpression as well as some other useful tools. By the end of this guide, you will be ready to start using the Xpression pipeline.

By convention, the virtual system is known as the 'guest', while the physical computer acts as the 'host'.

## Installation of Xpression VE

The only software Xpression VE needs to run is virtualization software such as Oracle's VirtualBox. This guide will use the freely available VirtualBox for demonstration, but others may be used.

If software like VirtualBox is not already installed on your computer, do so now. First browse to VirtualBox's website at https://www.virtualbox.org/wiki/Downloads and select the option for your operating system. Right-click the file and select 'save' or 'save-as' to download the file to your computer. After the download is finished, double-click this file to start the installation wizard. The other file you need to download is the Xpression RNAseq analysis pipeline.ova file from the Harwood lab website: https://depts.washington.edu/cshlab/html/rnaseq.html. Right-click on this file and save to your computer.

After a successful installation of VirtualBox, you can start VirtualBox. Once VirtualBox is started and the VirtualBox Manager window is open, you can import Xpression VE. Click on the file menu, then 'import appliance', and browse to find where the Xpression RNAseq pipeline.ova file was downloaded. Follow the instructions of this wizard to import the environment into VirtualBox. You may need to agree to the license of this file, but the other options can be modified at a later time.

Now you can see the Xpression environment in the VirtualBox Manager window. Only a few options should be changed before you can start Xpression.

- **Increase resources**
  The settings given as defaults are functional but, for decreased running time of Xpression as well as smoother operation of the other tools, it is recommended to increase the amount of resources given to the virtual machine.

  Click the 'Settings' icon to open the Settings window, and select the 'System' tab. The 'Motherboard' tab includes a slider-bar for 'Base Memory'. By entering a number or sliding this bar further to the right, you can increase the size of files and number of applications open at the same time in the virtual machine guest. The green area is considered 'safe', while any higher will slow the host system too much.

  Next, select the 'Processor' tab. If your hardware supports this option, a similar slider-bar for 'Processor(s)' can be modified as before. By sliding the bar further to the right while still in the green zone, the amount of computing power the virtual system is given is increased.

- **Link a folder to share**
  You should provide access to a folder on the host system. Using this folder for all of your activity within the virtual environment is recommended, as it is the most efficient way to access files without needless copying or transferring. Select the 'Shared Folders' tab in the left side of the Settings menu. This area allows you to choose which folder is shared between the virtual and physical computers. This folder is the only point of access the virtual computer has to your physical files.

  Click the  small folder button on the righ side of this window to add a folder. In the window that is opened, click the arrow on the 'Folder Path' bar and then 'Other...'. Browse to an appropriate folder on your physical computer. This folder could contain files like the sequencing file you wish to process, or could just be your home folder. This folder can also be used to store the results of Xpression. You cannot access the parent folder above the folder you linked to, or anywhere else in the file system that is not directly located within this folder.

  *Note: The files stored in this folder are your physical files and are not virtual in any way, where any modification including deletion is as real and permanent as in the physical computer.*

  Once this folder is chosen, change the Folder Name field to *shared*. This is very important, and failure to label this filed exactly *shared* will result in lack of access to this folder from within the virtual environment.

## Starting Xpression VE

Now that the settings have been modified for your particular computer, the Xpression virtual environment is ready to start. Select the Xpression item from the main window, then the start button in the upper-left corner. This will open a new window, which may be black for a while, or flash text, but this is normal. If any pop-up windows occur, select okay, since these are likely asking you to use the mouse cursor within this window as if it were a computer itself.

Once Xpression is fully started, you will see a file explorer window and icons on the desktop. The folder named 'shared' is located on the Desktop, and is the folder linked to your physical computer. This folder contains the actual contents of what is on your computer, and unlike the rest of the system, *is in no way virtual*. Whatever you choose to modify, copy or delete is just as permanent as when you do so from the physical computer.

## Interacting with the graphical environment

Although this environment may not be immediately familiar, it is quite similar to the Mac or Windows environment you may be more comfortable with. An application panel displays the time and which applications are running. Like the Windows start bar, you may switch to an application by clicking on it in the panel.

*Note: if you cannot see a grey bar at the bottom of the screen, your monitor's resolution may be too small. To see the entire window, press right-control (the virtualbox home key) and the 'c' key at the same time. This changes the window to 'scaling' mode, which will allow you to resize this window to suite your monitor's resolution while still seeing the complete virtual environment window.*

Right-clicking anywhere on the desktop or application panel will open a menu of applications to run. These applications include useful tools for interacting with the data from the Xpression pipeline. These include:

- **Xpression pipeline graphical user interface**
  This application analyses next-gen RNA-seq data, and is ready to use.
- **Integrated Genome Viewer**
  For viewing the visualized expression profiles resulting from data analysis by Xpression.
- **Internet browser**
- **Archive utility to compress or decompress files**
- **File explorer**
- **Spreadsheet editor**
- **Calculator**
- **Text editor**
- **Image viewer**
- **Xpression DEstat GUI**
  This tools uses the Bioconductor/R package DESeq to statistically compare two sets of RNA-seq data. Xpression generates a file of expression data for each sample analysed. For each condition in the comparision, provide one or more files, as well as a relavent label. Once each condition contains at least one file, click the 'compare' button. After processing is complete, two files will be created in the output folder chosen. The first file givens fold-change and p-values for each locus. The second is a MA-plot of the comparision.

Many of these tools will be opened automatically if you double-click a file as happens in MacOSX and Windows. For example, double-clicking 'Xpression GUI.jar' will start the Xpression application. By double-clicking the file of expression data that the Xpression pipeline generates, the spreadsheet editor will open this file. Likewise with an archive, you may either right-click and select 'Extract …' or double-click to open the archive utility.

Except for the Xpression GUI, these tools are not required by any part of the Xpression pipeline and are simply provided for convienience.

Now that your system is ready, please see the Xpression user guide for instructions on running the pipeline itself. For other guidance, please see the related Xpression documentation.