

Rare Variant Analysis Readme:

Hands-on Code: << Code.R or RareVariantsAnalysis.R >>

Hands-on Datasets : << UAB Rare Variants Demo.zip , Exercises.docx >>

Software programs demonstrated in the hands-on session youtube videos should be practiced while following the hands-on videos. The dataset files package of the course contents contains all the source codes and datasets required to carry out the analysis pipeline. For R-based analytics, one can either self-install the required libraries or use a pre-built virtual appliance which is preloaded with all the required tools to carry out the analytics pipeline.

Option A: To Self-Install the required libraries and packages:

Prerequisite Software/Tools/Libraries:

- R (3.3.1)
- R-Studio
- R-Packages: SKAT (1.12.1)

Above codes and packages are deployed and tested for R-3.3.1 at the time of preparing material for this on-line short-course.

Alternate to installing above mentioned R libraries and packages, one can benefit from our virtual machine(VM) appliance which is pre-loaded with all the libraries and packages required to perform the analysis modules.

Option B: To perform analysis on pre-build VM appliance, follow the steps:

Please visit the web site at <http://www.statgenlab.com/download/> for instructions on setting up your notebook/laptop computers to host the virtual appliance image. While we have tested the installation procedure on a number of different computers, there will be many different combinations of hardware and operating systems and some may require further troubleshooting. The end-to-end installation process should not take more than 20 minutes depending on your internet speed.

As on Nov 30th 2017, we recommend using Virtual Box release version 5.2.2.

In order to access “SKAT-Michael-Wu” within the virtual machine, please follow the below instructions:

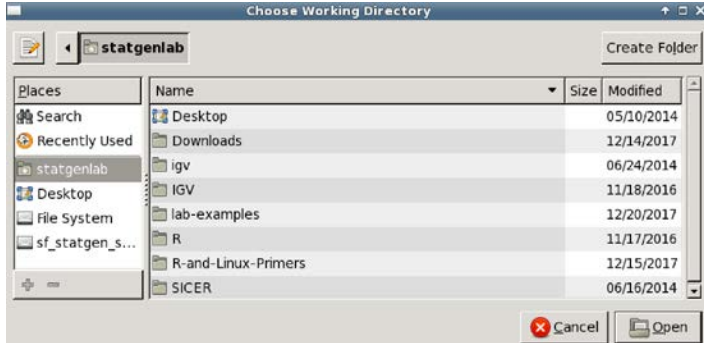
- Start the “StatGenLab” virtual appliance within Virtual Box.
- Start ‘R Studio’ application linked in the horizontal tray.



- Follow “Session -> Set Working Directory -> Choose Directory” menu options to select an analysis working directory.



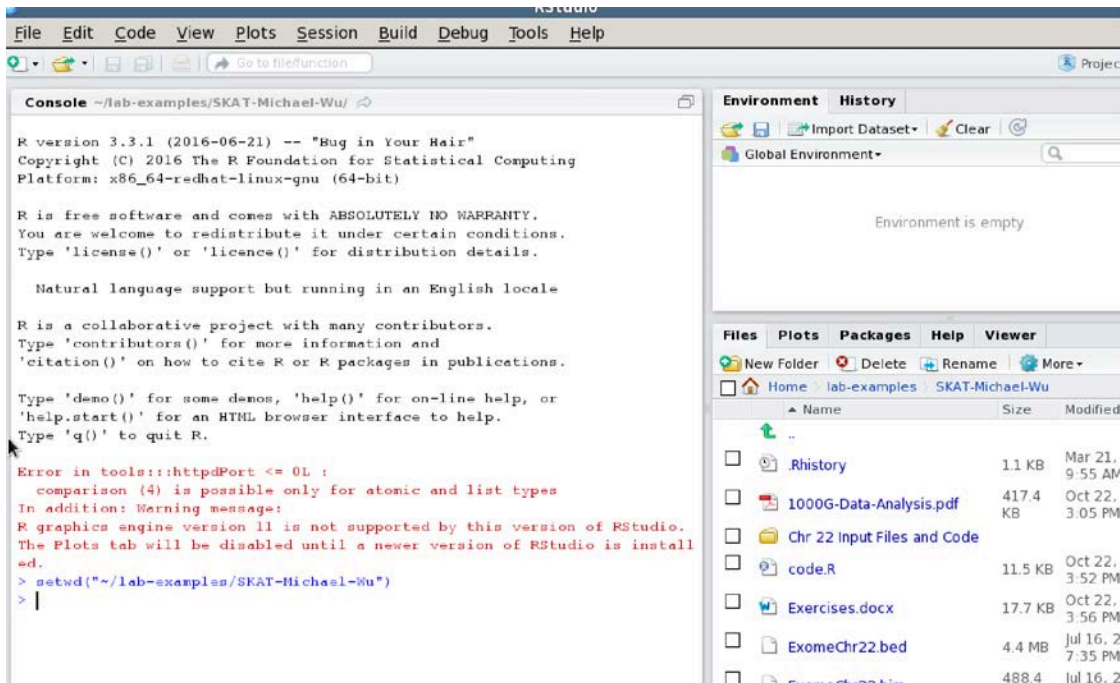
- Choose “lab-examples” folder.



- Within “lab-examples” folder, select “SKAT-Michael-Wu” folder.



- Upon selecting the analysis folder, R-Studio environment will list the files contained with that folder and will automatically set the current working directory to be the same.



- You may now follow the Youtube instructions for this analyzing this datasets package.
- << Code.txt>> filename needs to be changed to <<Code.R>>. This source-code is same as << RareVariantsAnalysis.R>>.