VisFeature: a stand-alone program for visualizing and analyzing statistical features of biological sequences

A User Manual

Contact	Jun Wang wj0708@tju.edu.cn
Date	August 8, 2019
Version	1.0

Contents

1 Introduction

In order to predict biological functions and cellular attributes, one of the most challenging problems is to find a valid approach to visually represent the features of biological sequences. Therefore, it is important to analyze the features of biological sequences. We developed VisFeature, a stand-alone program for visualizing and analyzing statistical features of biological sequences. It provides a function for retrieving a number of biological sequences automatically and integrates over 30 sequence representation modes for extracting the features of DNA, RNA and protein sequences rapidly. It can also convert sequences into curves and density maps based on their physicochemical properties or feature vectors. By using the feature visualization, the differences between different sequences or groups can be observed, which is convenient for sequence analysis.

There are four main functions in VisFeature: "Fetch data", "Single sequence mode", "Multiple sequences mode" and "Density map comparison".

You can use "Fetch data" to download a large number of DNA, RNA and protein sequences by identifier or query expression. You can use "Single sequence mode" to visualize a DNA, RNA or protein sequence by physicochemical properties. You can use "Multiple sequences mode" to visualize multiple DNA, RNA or protein sequences by physicochemical properties. You can also perform multiple sequence alignment in this function. You can use "Density map comparison" to calculate feature vectors from DNA, RNA and protein sequences. After the calculation, a label file should be uploaded to assign group labels to each sequence. You can use this function to visualizing and comparing feature vectors as density maps.

2 Usage

2.1 Instructions for MacOS users

VisFeature is a software for Microsoft Windows. It is not designed to be a cross-platform software. We provide a Linux version only for the linux users to experience the basic functions of VisFeature. The only recommended version is the Windows version.

For **Mac OS users**, we provide a solution to experience the basic functions of VisFeature using a Linux Virtual Machine. See below instructions for details.

2.2 Install VirtualBox Hypervisor

Use safari to access the https://www.virtualbox.org and download the VirtualBox for OS X hosts.
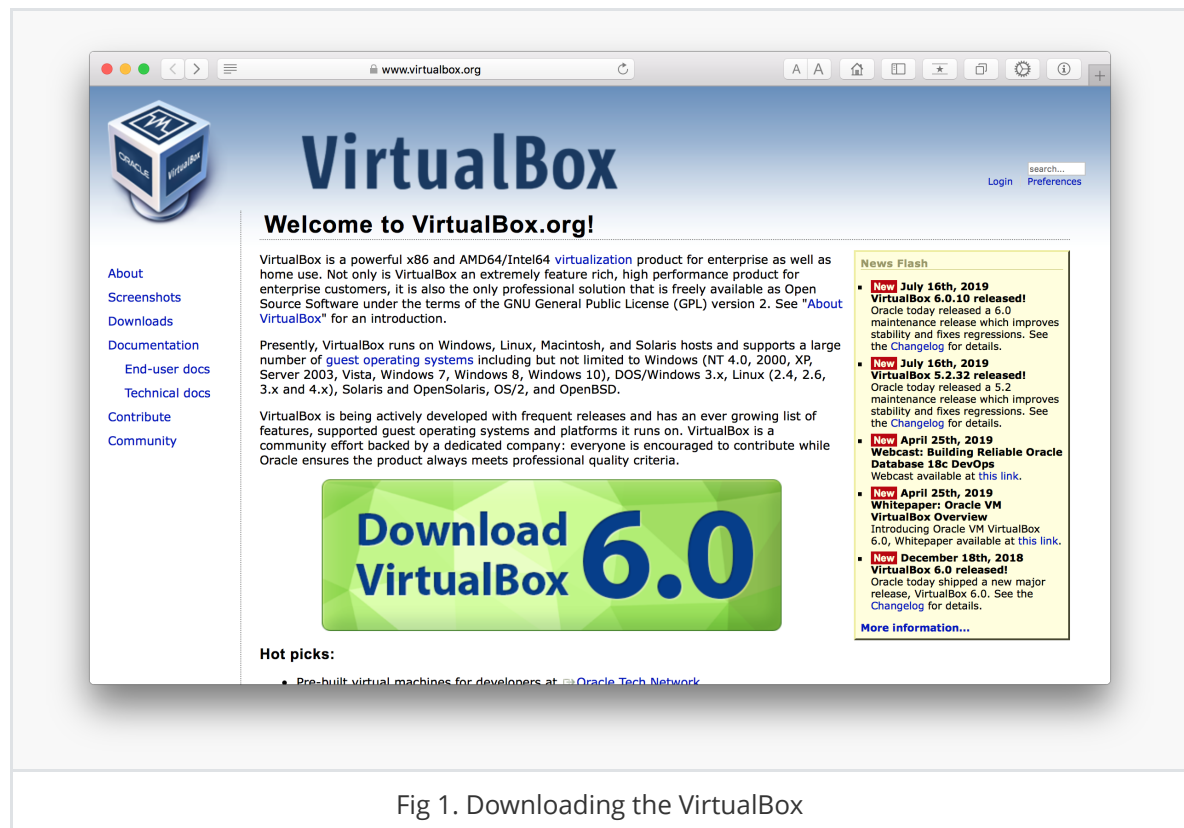

Fig 1. Downloading the VirtualBox

Choose the OSX hosts version VirtualBox

Fig 2. Choose the OSX hosts version

Install the VirtualBox as a normal application on your Mac.

After the installation of VirtualBox, execute the VirtualBox application.

Fig 3. Start VirtualBox

2.3 Import the VisFeature Ubuntu image

Download the `VisFeature-macOS.zip` from https://github.com/wangjun1996/VisFeature/releases. Unpack it to obtain an .ova file. This file is the VisFeature Ubuntu image.

Click the "Import" button on the VirtualBox main interface. Fill in the location of the VisFeature Ubuntu image.
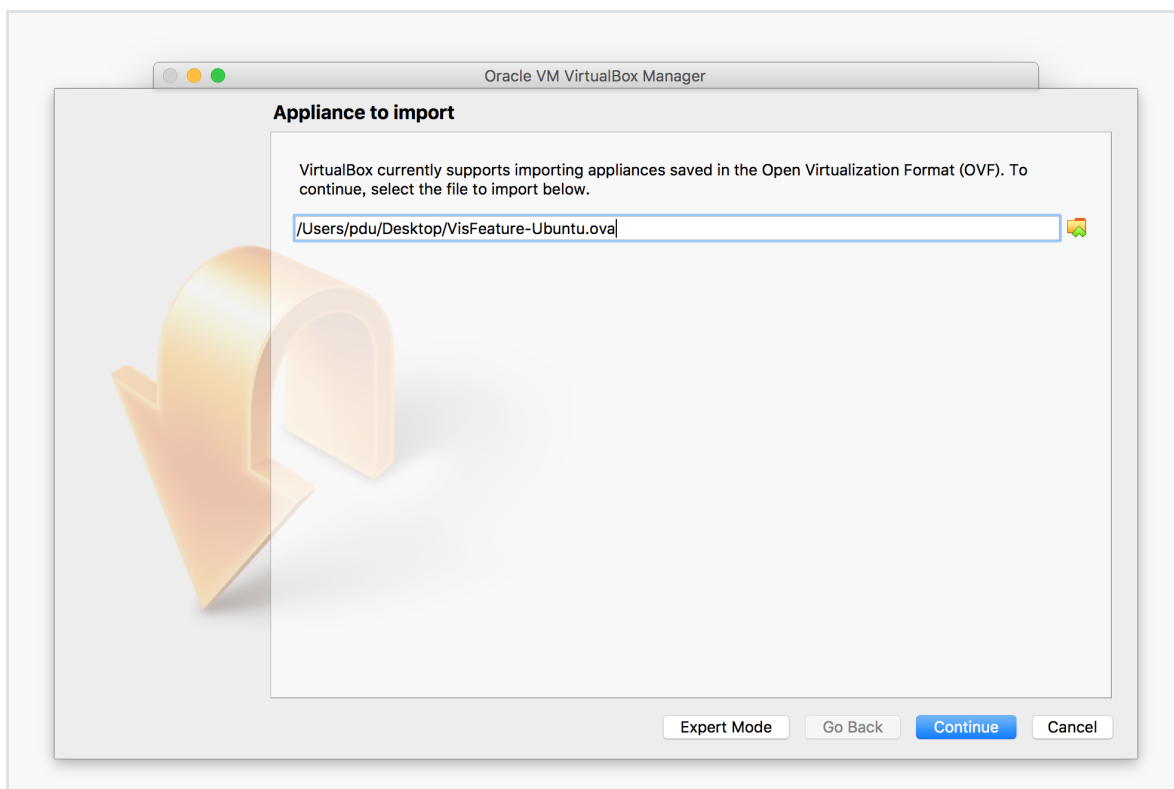
Fig 4. Import Ubuntu Image

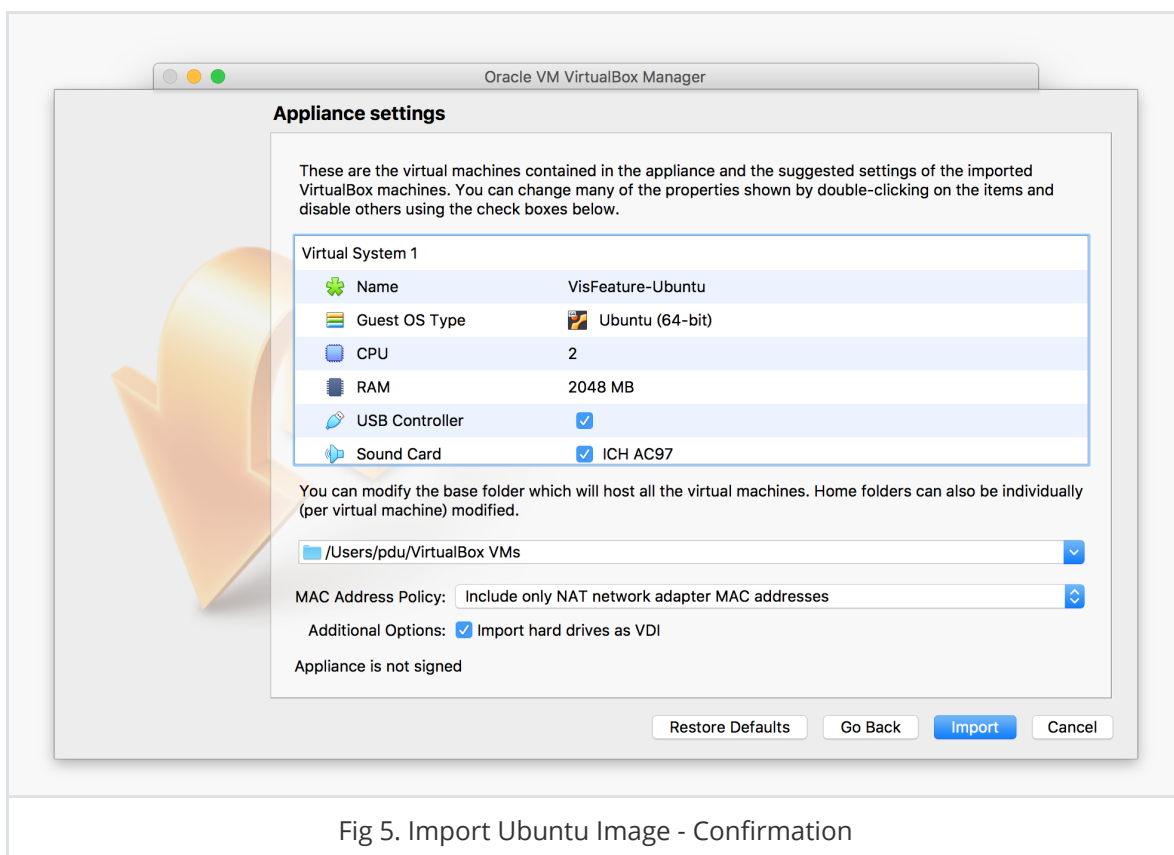Click "Continue" to the next page.

Fig 5. Import Ubuntu Image - Confirmation

Make sure you have sufficient processors and RAM on your host machine. The VisFeature-Ubuntu image require **at lease 2 cores and 2GB memory** to be operable. Your physical machine should have at **least 4 cores and 4GB memory** to load this image. In addition, your physical processor MUST support **VT-x or AMD-V** to load this image.
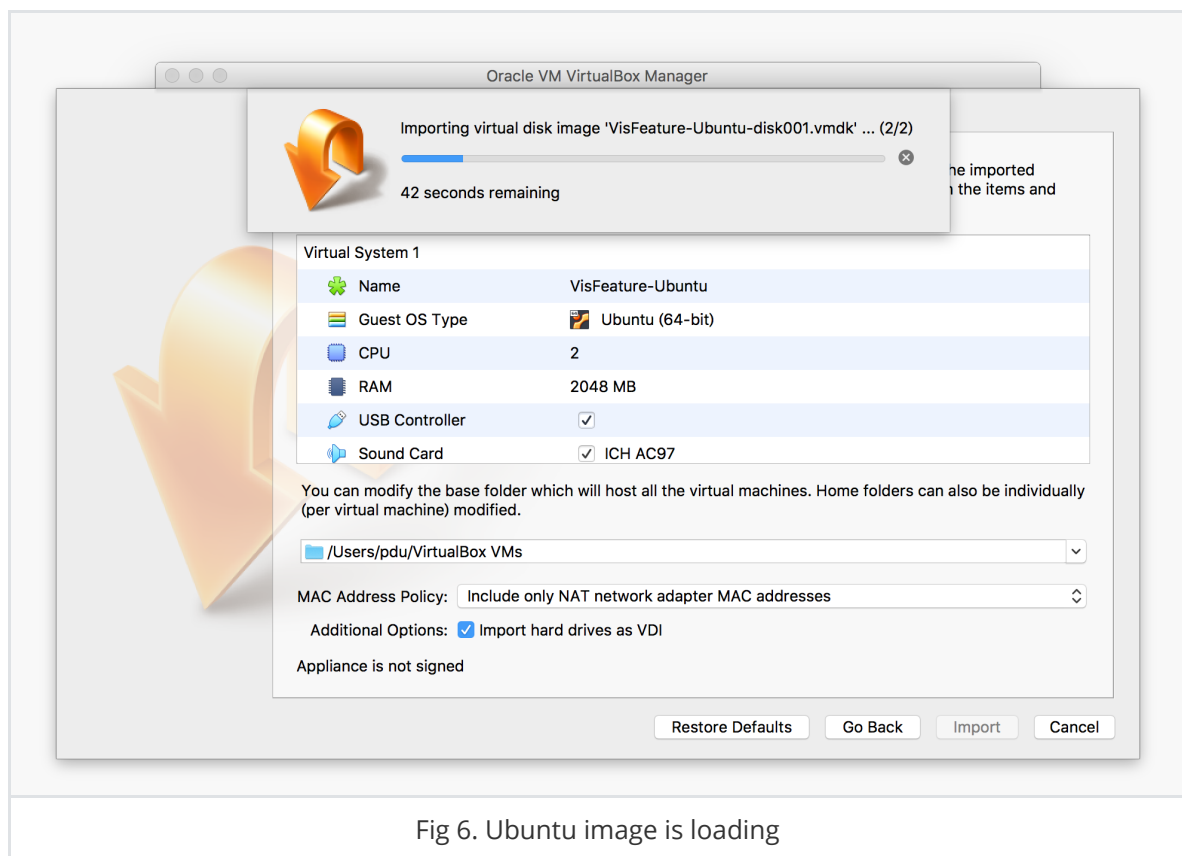
Click "Import" to load the image.

Fig 6. Ubuntu image is loading

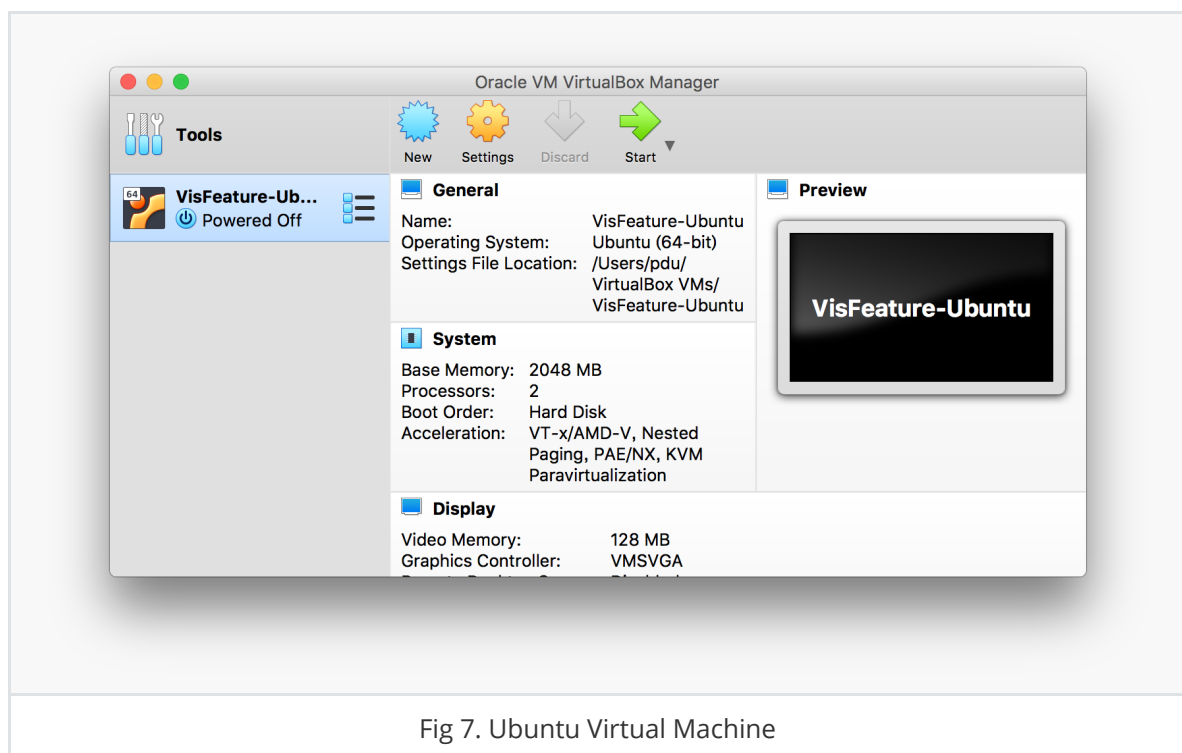After the loading process, you will see a Ubuntu VM in your VirtualBox.

Fig 7. Ubuntu Virtual Machine

2.4 Use VisFeature in the Virtual Machine

Click "Start" to start the Ubuntu virtual machine. You will be asked if you would like it to be in full screen mode. You may read the instructions and then click "Switch".
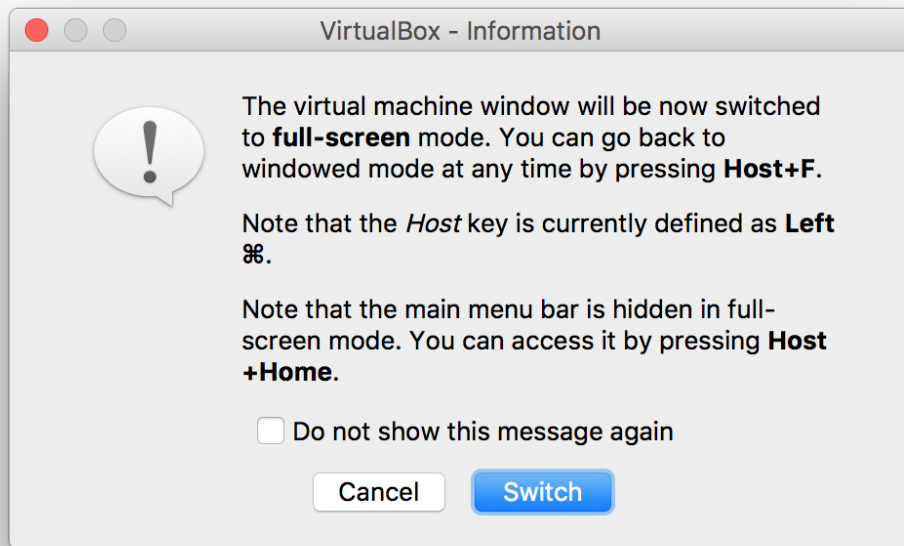
Fig 8. Switch to full screen

After several seconds, you will be directed to a Ubuntu desktop as below. The VisFeature and all the dependencies have been pre-loaded. Just double-click the "VisFeature" icon on the desktop. **The remaining operations are identical to the Linux versions**.
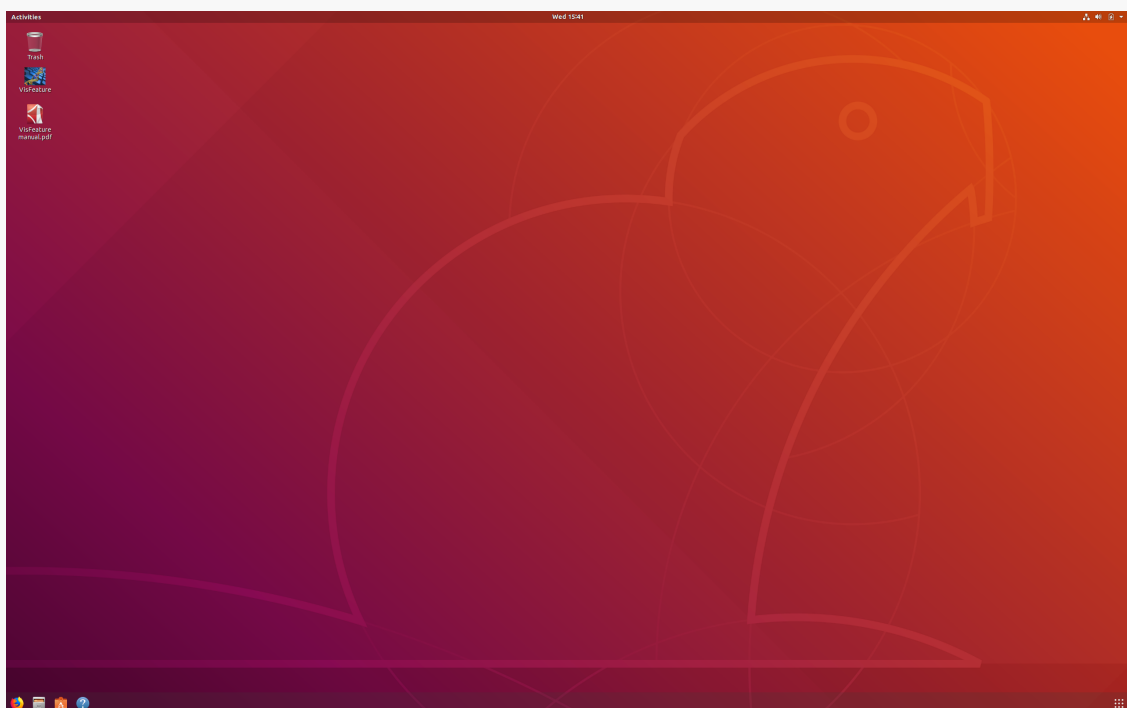

Fig 9. Ubuntu image Desktop in VirtualBox

Please read the manual for Linux versions from https://github.com/wangjun1996/VisFeature/tree/master/doc for detailed operations. Please also read the manual of VirtualBox on the https://www.virtualbox.org to understand the basic operations and data exchange between VirtualBox VMs and the host machines.

3 Other Information

3.1 Development

If you want to run these codes in the development environment, you should install **Node.js** first. Please go to https://nodejs.org/en/download/ download Node.js environment and install that on your machine.

After your Node.js environment is ready, find out the location of the source code of VisFeature that you unpack and enter this directory in command line program. Then type and execute the command: `npm start`. After a few seconds, VisFeature will start.

If you want to **package** application, you should install **electron** and **electron-packager** additionally by executing the command `npm install electron -g` and `npm install electron-packager -g` in command line program. Then, find out the location of the source code of VisFeature that you unpack and enter this directory. Finally, type and execute the command: `npm run windows` or `npm run linux` to get corresponding binary release.

3.2 Authors correspondence

If you have any questions or bug reports about VisFeature, please feel free to contact JunWang, Email: wj0708@tju.edu.cn.

3.3 Submitting your add-ons

If you want to contribute plugins for VisFeature, just join the project on GitHub. The address of VisFeature is https://github.com/wangjun1996/VisFeature.

3.4 Resources

The following resources may be useful when you use VisFeature.

- UltraPse

https://github.com/pufengdu/UltraPse

- How to fetch data by Uniprot API

https://www.uniprot.org/help/api

- How to fetch data by NCBI API

https://www.ncbi.nlm.nih.gov/home/develop/api/

- Clustalw2

https://www.ebi.ac.uk/Tools/msa/clustalw2/

- Electron

https://electronjs.org/

- Echarts

https://github.com/apache/incubator-echarts

- R

https://www.r-project.org/

References

- Li,W.-C. et al. (2015) iORI-PseKNC: A predictor for identifying origin of replication with pseudo k-tuple nucleotide composition. Chemometrics and Intelligent Laboratory Systems, 141, 100–106.