

VICTOR provides four modes of usage. All requires that the usage of any contents of this software package is for research purpose only.

A. Install to your local machine

Install

1. Make sure your computer has the supported CPU architecture (X86_64) and operating system (Linux kernel 2.6.18 or later, or Mac OS 10.7 or later).
2. Decide where to put VICTOR (for example: /path/to/VICTOR/). Replace all the "/path/to/" below with the desired location.
3. Install programs.
 - 1) go to the install location (cd /path/to/),
 - 2) download the program package for Linux (wget https://www.dropbox.com/s/rkfu47k8o4mh5wg/VICTOR_linux.tgz) or download the program package for Mac OS (wget https://www.dropbox.com/s/shu32s3r0rhv5j8/VICTOR_mac.tgz),
 - 3) uncompress ("tar xzvf VICTOR_linux.tgz" or "tar xzvf VICTOR_mac.tgz"),
 - 4) add the location to your \$PATH (add the line "PATH=/path/to/VICTOR:\$PATH" to your .profile file), or create a module file with the line "prepend-path PATH /path/to/VICTOR" and load the module before each usage.
4. Install data. Go to VICTOR's data directory (cd /path/to/VICTOR/data/) and execute one of the following:


```
curl -sL http://www.bjfenlab.org/download/get_hg19_inst.sh | bash
curl -sL http://www.bjfenlab.org/download/get_GRCh37_inst.sh | bash
curl -sL http://www.bjfenlab.org/download/get_GRCh38_inst.sh | bash
```

 They will download and uncompress the data files to the current directory.
5. Install the latest data updates. Some data, such as ClinVar, need to be updated more frequently. I put them in a monthly update file that has a version number different from the data folder. So, even your data folder is up-to-date, you may still need to install the latest updates. Go to VICTOR's data directory (cd /path/to/VICTOR/data/) and do one the following:


```
curl -sL http://www.bjfenlab.org/download/get_hg19_update.sh | bash
curl -sL http://www.bjfenlab.org/download/get_GRCh37_update.sh | bash
curl -sL http://www.bjfenlab.org/download/get_GRCh38_update.sh | bash
```
6. Install R packages. Run the following commands within R:


```
install.packages("logistf")
```
7. Install third-party programs. Below is a list of the programs. The minimum requirement is tabix and GNU parallel. PROVEAN, PLINK and KING are highly recommended. Make sure these programs are in your \$PATH.

Programs	Comment	It will be used for	License	URL
tabix	required	retrieval of file contents	MIT	https://sourceforge.net/projects/samtools/files/tabix/
GNU parallel	required	parallel computing	GPLv3	https://www.gnu.org/software/parallel/
PROVEAN	recommended	PROVEAN score calculation	GPLv3	http://provean.jcvi.org/
blast	recommended	PROVEAN	Public	https://blast.ncbi.nlm.nih.gov/
CD-HIT 4.5.8	recommended	PROVEAN	GPLv2	http://weizhongli-lab.org/cd-hit/
PLINK 1.9	recommended	quality control	GPLv3	https://www.cog-genomics.org/plink2/
KING	recommended	quality control	Unknown	http://people.virginia.edu/~wc9c/KING/
ShapeIt2	optional	phasing	Academic	https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html
BEAGLE	optional	phasing	GPLv3	https://faculty.washington.edu/browning/beagle/beagle.html
GATK	optional	VCF file combining	BSD	https://software.broadinstitute.org/gatk/download/
gnuplot	optional	drawing Manhattan plots	gnuplot	http://www.gnuplot.info/

I have made a bundle file with the above software except GNU parallel and gnuplot to help you install them altogether. Go to the installation directory ("cd ~/local/" as an example) and do the following. This command will create a folder named "pkg" and put all programs and databases in it. Afterwards, you need to add PATH in the .bash_profile file or create a module file to set PATH. License of each software is within the bundle.

```
curl -sL http://www.bjfenlab.org/download/get_pkg_inst.sh | bash
### please add "PATH=$HOME/local/pkg:$HOME/local/pkg/PROVEAN/bin:$PATH" to ~/.bash_profile, or
### make a module file with "prepend-path PATH $HOME/local/pkg" and "prepend-path PATH $HOME/local/pkg/PROVEAN/bin".
```

Upgrade

To check whether your programs or data are up-to-date, you can type "vQC --genome=GRCh37 --version". Change the --genome option if you want to check other genome data. If you type this command within a directory whose fullpath contains a genome name, this option can be omitted. Not only vQC, all other programs support the --version option and will give the same results. If the data is up-to-date but the programs are not, you only need to upgrade the programs. The procedure to upgrade is the same as installation.

After each upgrade, please re-make the slurm.all_steps script again with the new template if you want to re-analyze the data using the newer version of VICTOR, because it is likely that there are important changes in the template.

Monthly data updates

Please see the above bullet point number 5.

B. Do analyses online (not provided in the above downloadable package)

Pedigree maker

This server helps to make a pedigree drawing and creates a pedigree file in LINKAGE format (LINKAGE is a program for linkage analysis; this format is widely used in genetic analysis including variant classification by VICTOR). It starts from a proband or a pedigree file and recursively adds a relative to any family member specified by the user. The relative could be a distant one. For example, you can add an uncle to the proband even when the parents and grand-parents of the proband are not yet in the pedigree. The program will automatically add these individuals with missing affection status and age. The missing values can be changed later. This makes pedigree creation fast when you don't need to add information for all family members.

When the pedigree input is complete, you can right-click on the image and select "view frame source" or "save frame as", which should save a .svg file that could be opened by most web browsers and print out a PDF or other image files. It can also output a pedigree file, which is compatible with the variant classifier program below. If you have multiple pedigrees that carry the same variant, do this process several times and concatenate the files into one (please keep only one header line at the top of the file) and upload it to the variant classifier below.

Features:

- * Input from a pedigree file or manually create a pedigree;
- * Make a pedigree drawing;
- * Output a new pedigree file;
- * Automatically check for pedigree file errors and correct minor mistakes (e.g. wrong sex, missing individuals);
- * The input pedigree file format is flexible.
- * Compatible with the Variant Classification program below.
- * Allow multiple diseases caused by one gene (e.g., BrCa and OvCa by *BRCA*).

Quantitative variant classification

This server calculates BayesDel (Feng, Hum Mut, 2017), performs co-segregation analysis (Thompson et al., AJHG 2003), and annotates allele frequency and ClinVar classification.

Features:

- * It handles missense substitution at the same position as a ClinVar variant(s) but to a different amino acid.
- * If a protein-level HGVS have multiple genomic interpretations, the program will automatically select the most probable one.
- * It provides built-in penetrance models for BRCA1 and BRCA2. It can work for other genes if a model is provided.
- * If a proband is not provided, it will automatically choose a proband in a conservative way.
- * It uses a Mutation Rate parameter to analyze genes with high de novo mutation rates, such as TP53.
- * Allele frequency is obtained from gnomAD, UK10K, GoNL, and 1kJpn.

For co-segregation analysis, it takes a pedigree file. Here are examples of the pedigree file in [BOADICEA format](#) and [LINKAGE format before "makeped"](#) (10 columns for PedID, IndivID, FathID, MothID, Sex, AffectionStatus, Liability, Proband, Allele1, Allele2) and [LINKAGE format after "makeped"](#).

C. Amazon cloud computing

Amazon Web Service (AWS) is a HIPAA-compliant hosting and cloud computing provider. The Elastic Compute Cloud (EC2) is one of the services AWS has to provide. I have installed VICTOR and all third-party programs in an EC2 instance and created an Amazon Machine Image (AMI). You can launch an EC2 instance with your AWS account using this image and do analyses on the cloud. The image is free for you, but any cost for the computing and storage is between you and Amazon. I don't get any rewards from Amazon for providing this image.

To do cloud computing, in the AWS Management Console (<https://aws.amazon.com/console/>) choose "Instances", then "Launch", click "Community AMIs" and search for "victor_GRCh37". Select the image and go on with the instance launching process. You need to choose a storage capacity that can host the operating system (6G), VICTOR (14G), other programs (13G), and your data and result files (30G or more) based on the size of your study. So the total storage may be 65G. After launching, you can enter the directory \$HOME/GRCh37, upload your data, copy slurm.all_steps to the folder and set parameters (you may need to change PRL depending on the number of CPUs), directly run the script following the instructions inside the script, get results, then terminate the instance.

D. Source code

The c++ source codes of VICTOR programs are at github (<https://github.com/fengbjlab/victor/>).