

$\Delta_{\text{vina}}\text{RF}_{20}$ Tutorial

Cheng Wang

2016

First Part of this tutorial is about installation of prerequisites and set up the DeltaVina code. Second Part is an example of applying $\Delta_{\text{vina}}\text{RF}_{20}$ scoring function to rescore the protein-ligand binding affinity.

DeltaVina code

It can be accessed from github.

Github: <https://github.com/chengwang88/deltavina>

1. Prerequisites

- MGLTools and MSMS
- A modified version of AutoDock Vina
- Python with pandas and openbabel libraries
- R with randomForest library
- Set the environment variable

1.1 MGLTools and MSMS

MGLTools are used for preparing PDBQT file for Vina. MSMS is for calculating solvent accessible surface area. Both of them can be downloaded from MGLTools website.

Web: <http://mgltools.scripps.edu/downloads>

Notes:

- MGLTools can be installed by

```
tar -xvzf mgltools_x86_64Linux2_1.5.6.tar.gz
cd mgltools_x86_64Linux2_1.5.6/
./install.sh
```

- After downloaded msms_i86_64Linux2_2.6.1.tar.gz

```
mkdir msms
tar -xvzf msms_i86_64Linux2_2.6.1.tar.gz -C msms
cd msms
```

- Copy `msms.x86_64Linux2.2.6.1` to be `msms`

```
cp msms.x86_64Linux2.2.6.1 msms
```

- In `msms` folder, there is a script `pdb_to_xyzr`. Change the line
`numfile = "./atmtypenumbers"`
to be
`numfile = "YourPATHtoMSMS/atmtypenumbers"`

1.2 A modified version of AutoDock Vina

A fork of Autodock Vina has been modified to output the features in `score_only` mode.

Github: <https://github.com/chengwang88/vina4dv>

1.3 Python with pandas and openbabel libraries

Python 2.7 from Anaconda is recommended since the openbabel python binding can be directly installed.

Web: <https://www.continuum.io/downloads>

Notes

- If you are the first time to use Anaconda python, source the `.bashrc` (Linux) or `.bash_profile` (macOS) may be needed in order to use conda.

```
source ~/.bashrc
conda update conda
```

- Install pandas

```
conda install pandas
```

- Install OpenBabel from <https://anaconda.org/RMG/openbabel>

```
conda install -c rmg openbabel=2.3.2
```

1.4 R with randomForest library

Most recent version of R is recommended.

Web: <https://cran.r-project.org/>

Note

- Install randomForest in R by

```
install.packages('randomForest')
```

1.5 Set the environment variable

If you have the prerequisites installed already and DeltaVina downloaded from github. Several environment variables need to be set in `.bashrc` (Linux) or `.bash_profile` (macOS) file in your home directory. An example is given below. You can modify the path based on your case. In this example, all softwares are installed under `/home/vagrant/sw` directory.

Example of environment variable setting

```
# path for MSMS
export PATH=$PATH:/home/vagrant/sw/msms/

# set mgltool variable
export MGL=/home/vagrant/sw/mgltools_x86_64Linux2_1.5.6/
# set mgltool python
export MGLPY=$MGL/bin/python
# set mgl utilities path
export MGLUTIL=$MGL/MGLToolsPckgs/AutoDockTools/Utilities24/

# path for deltevina
export PATH=/home/vagrant/sw/deltavina/bin:$PATH
# pythonpath for deltevina
export PYTHONPATH='/home/vagrant/sw/deltavina'

# set vina dir
export VINADIR=/home/vagrant/sw/vina4dv/build/linux/release/
```

2. Example of using DeltaVina

After all of above have been set up, the example can be run in `deltavina/examples`.

```
cd examples
```

You can check the help by

```
dvr20.py --help
```

The script can be run for one complex by

```
dvr20.py -r 1a42/1a42_protein_proc_se.pdb -l  
1a42/1a42_ligand_fix.mol2
```

Or it can also be run by providing a list of protein-ligand complex with input format as
in `pdblist.txt`

```
cat pdblist.txt  
1a42/1a42_protein_proc_se.pdb 1a42/1a42_ligand_fix.mol2  
dvr20.py -g pdblist.txt
```