

rvd_gen User Manual

Background Information

rvd_gen is a command line tool that can be used to generate a designer TAL effector RVD sequence from a given nucleotide sequence. Because rvd_gen is written in C, it should run on any modern UNIX-like operating system with little to no modification. It has been tested and verified as functional on GNU/Linux, macOS, FreeBSD, and Windows (via Cygwin). For the rvd_gen software license and terms, see the end of this document.

Installation Instructions

GNU/Linux, macOS, or similar command line operating systems

rvd_gen can be installed by cloning from github using the **git** command line tool (for installation of git, see the developer website at <https://git-scm.com/downloads>). To install rvd_gen, use **git** to clone the distribution, then use **make** to compile, as below:

```
~$ git clone https://github.com/stephen-cohen/rvd_gen
Cloning into 'rvd_gen'...
Checking connectivity... done.
~$ cd rvd_gen
~/rvd_gen$ make
cc -Os -Wall -o rvd_gen rvd_gen.c
```

The output binary file is **rvd_gen**, which can be run directly or copied to a directory in \$PATH (e.g. ~/bin for a single user, or /usr/local/bin in Linux for multiple users) for easy access.

Windows instructions

rvd_gen can be installed in Windows using GNU tools in Cygwin (downloadable at <https://www.cygwin.com/>). When installing Cygwin, install the development tools **make**, **gcc-core**, and **git**. Using the **git** and **make** commands as described above will produce **rvd_gen.exe**.

Usage Instructions

To use rvd_gen, you must supply a nucleotide sequence. The sequence may be supplied as a command line argument or as input, as below:

```
~/rvd_gen$ ./rvd_gen ACGTACGTACGTACGT
NI HD NK NG NI HD NK NG NI HD NK NG NI HD NK NG
~/rvd_gen$ ./rvd_gen
ACGTACGTACGTACGT
NI HD NK NG NI HD NK NG NI HD NK NG NI HD NK NG
```

By default, the maximum sequence length allowed is 41, because most available dTALE construction kits do not allow target sequences to be larger than ~30. Source code (rvd_gen.c, line 41) may be altered to raise or lower this limit. There is no minimum input for the program. Input may be upper- and/or lowercase.

About the Program

This program generates an optimal TAL effector RVD sequence given a nucleotide sequence. Table 1 shows the nucleotide-RVD pairs used by `rvd_gen`, which are available in modules from the dTALE construction kit made available by Cermak et al. (2011)¹.

Table 1. Nucleotide-RVD pairs used by `rvd_gen` and the relative strength of each corresponding interaction.

Nucleotide(s)	RVD	Strength ²
A	NI	weak
C	HD	strong
G	NH / NK	weak
T	NG	weak
A / G	NN	strong

In the case of sequences that would lead to stretches of 6 or more weak RVDs, `rvd_gen` substitutes NN for binding to A or G to break up the weak stretches, based on recommendations from Streubel et al. (2012)². Input sequences with stretches of 6 or more Ts cannot be strengthened and should be avoided as EBEs. `rvd_gen` outputs NK for G specificity, but NH may be substituted.

A user may force usage of the RVD NN in a specific location by using the nucleotide symbol R, as below:

```
~/rvd_gen$ ./rvd_gen AAAAAA
NI NI NI NI NI NN
~/rvd_gen$ ./rvd_gen ARAAAA
NI NN NI NI NI NI
```

Uninterpretable input returns “??” in place of valid RVDs.

References

1. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF. Nucleic Acids Res. 2011; 39(12):e82.
2. Streubel J, Blücher C, Landgraf A, Boch J. Nat Biotechnol. 2012;30:593-595.

Citing `rvd_gen`

To cite `rvd_gen` in publications, use:

Cohen SP. 2018. `rvd_gen`: a command line tool for generating designer TAL effector RVD sequences. Available from: https://github.com/stephen-cohen/rvd_gen.

Software License

Copyright © 2017-2018, Stephen P. Cohen
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- Neither the name of nor the names of its contributors may be used endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.