

US EPA TOXCAST DATA RELEASE SEPTEMBER 2019 Database & tcpl R Package

This file describes the contents of the September 2019 ToxCast data release. The database/R package contents of the download contain the following files:

Invitrodb Database Loading

To utilize the database, the user must have write access to a MySQL server, allowing them to install the database contained in the “prod_external_invitrodb_v3_2.all.zip” file. The “prod_external_invitrodb_v3_2.all.zip” is a zipped sql file containing the entire invitrodb database, whereas the “prod_external_invitrodb_v3_2.bare.sql” contains just the database schema and tcpl methods pre-loaded. To avoid errors while calling large queries we recommend the MySQL server setting, ‘max_allowed_packet’, is set to at least 1 GB.

Once you have write access to a MySQL server you can load the database. First, unzip the invitrodb-loader.zip and prod_external_invitrodb_v3_2 files to the same location as prod_external_invitrodb_v3_2.bar.sql. Next, follow the instructions for your operating system below to write to your database.

Windows - Double-click Load-InvitroDB.bat, then enter your server hostname, username, port number, and password, as prompted

Mac or Linux - Run ./Load-InvitroDB.sh in a terminal from the directory where you downloaded and unzipped the file

TCPL package installation

To install the package on Windows the user will first need to properly install R tools, enabling R to compile the source package. The package relies on R ($\geq 3.1.0$), and the following packages: data.table ($\geq 1.9.4$), DBI, MASS, numDeriv, RColorBrewer. To interact with the MySQL database, the user also needs the “RMySQL” package. The “RMySQL” package requires a local MySQL server, although it does not require the server to be running. More information about installing RMySQL on Windows can be found online.

The “tcpl_2.0.tar.gz” file is the beta version of the R package used to process all of the data in this release. The package can be installed using the following R command:

```
install.packages("tcpl_2.0.0.tar.gz", type = "source", repos = NULL)
```

The html files are the R package vignette, detailing the processing, and showing some examples of how to use the package. The appendices vignette also includes an appendix listing the database field names.

For questions or concerns, please contact Jason Brown at: brown.jason@epa.gov.