The RVDB (Reference Virus DataBase) is a collection of all known viral, viral-related, and viral-like sequences, except bacterial viruses. The RVDB is updated directly from the NCBI RefSeq, NCBI GenBank and NCBI TPA sequences collections. The files contained in this directory contain instructions for performing an update of the RVDB, as well as characterizing the resulting sequences by the following categories: exogenous viral, endogenous nonretroviral, endogenous retroviral, LTR-retrotransposon, and unassigned viral genes / fragments. Further instructions for performing the RVDB update are provided in the Supplementary Methods S1 document, while instructions for characterization are provided in the Supplementary Methods S2 document. The “RVDB\_characterization.py” python script is used for the characterization; all other python scripts are used for updating RVDB. The manuscript describing RVDB is currently under review (Norman Goodacre, Aisha Aljanahi, Subhiksha Nandakumar, Mike Mikailov, Arifa S. Khan. A reference viral database (RVDB) to enhance bioinformatics analysis of high throughput sequencing for novel virus detection).