

NLM Citation: BLAST[®] Command Line Applications User Manual [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2008-. Clean-up unused database volumes. 2023 Jun 30 [Updated 2023 Jun 30].

Bookshelf URL: https://www.ncbi.nlm.nih.gov/books/



Clean-up unused database volumes

Created: June 30, 2023; Updated: June 30, 2023.

The script cleanup-blastdb-volumes.py can be used to remove unused BLAST database volumes that may go unnoticed and take up space. There are two main ways that this may occur:

- The number of volumes in a BLAST database grew larger than 100, and the volumes have three digits rather than two (e.g., nr.02 becomes nr.002)
- The number of volumes in a BLAST database was reduced. This may happen if enough sequences are removed from the database or if the database was recreated with a larger maximum volume size (e.g.: via makeblastdb 's -max_file_sz command line option).

You may invoke the script with a simple command. For example, to remove unneeded volumes of nt, use the command below. For other databases, simply replace nt by the name of the database. For protein databases, replace "nucl" by "prot".

\$ cleanup-blastdb-volumes.py -db nt -dbtype nucl

cleanup-blastdb-volumes.py is included in the BLAST+ release.