Datasets in Computational Biology

Overview of Popular Datasets

Examples include GenBank, TCGA (The Cancer Genome Atlas), and PDB (Protein Data Bank).

Data Collection Methods and Challenges

Challenges include heterogeneity of data formats, large volumes, and data curation needs.

Comparative Analysis of Datasets

PDB is rich in 3D structural data, while TCGA focuses on cancer-related genomic information.