Table 1 Large and commonly used chemoinformatic resources

Name	Bioactivities	Link	Summary	Last update
ChEMBL	Various bioactivities (K_i, EC50 …), pChEMBL	https://www.ebi.ac.uk/chembl/	~1.6M compounds, 14M bioactivities, 11K targets	2016
K_i DB	K_i	http://kidbdev.med.unc.edu/databases/kidb.php	~10K compounds, 60K bioactivities, 746 targets	2016
BindingDB	Various bioactivities	https://www.bindingdb.org/bind/index.jsp	~552K compounds, 1.2M bioactivities, 6.4K targets	2016
PharmGKB	Drug response data	https://www.pharmgkb.org/	~27K genes, 3.5K diseases, 3.2K drugs, 108 pathways	2016
Guide to Pharmacology	Various bioactivities	http://www.guidetopharmacology.org/	~8.6K compounds, 15K bioactivities, 2.8K targets	2016
DrugBank	Drug–target interactions	http://www.drugbank.ca/	~8.2K drugs, 15K drug–target interactions, 4.2K targets	2016
CTD	Chemical–gene interactions, gene–disease and chemical–disease associations	http://www.ctdbase.org/	~1.4M chemical–gene interactions, 20M gene–disease associations, 2.1M chemical–disease associations	2016
STITCH	Association scores	http://stitch.embl.de/,	STITCH 5: ~9.6M proteins, 932M interactions,	2016

		http://stitch-beta.embl.de/	2K organisms	
PubChem	Various bioactivities	https://pubchem.ncbi.nlm.nih.gov/	~2.3M tested compounds, 231M bioactivities, 10K targets	2016
PHAROS	Various bioactivities, target–disease score	https://pharos.nih.gov/	~163K compounds, 166K bioactivities, 20K targets, 8.7K diseases	2016
Open Targets	Target–disease and drug–target associations	https://www.targetvalidation.org/	~2.4M target–disease associations, 9.4K diseases, 30K targets	2016
DGIdb	Drug–gene interactions	http://dgidb.genome.wustl.edu/	Without PharmGKB: ~12K compounds, 26K structure–gene pairs, 3.2K targets	2016
CARLSBAD	CARLSBAD activity	http://carlsbad.health.unm.edu/	~435K structures, 933K bioactivities, 3.7K targets	2014
ChemProt	ChemProt activity	http://potentia.cbs.dtu.dk/ChemProt/	~1.7M structures, 7.8M bioactivities, 19K targets	2016

The databases listed collect drug–target or drug–gene interactions mainly from published studies and from other databases. Some of them provide various drug–target bioactivities (K_i, IC50, etc.); others produce their own unique activity from experimental measures (pChEMBL, CARLSBAD activity, ChemProt

activity) or only indicate whether or not there is a drug–gene interaction (for example, DGIdb) and, if so, the type of interaction. Databases providing a collection of bioactivities (such as ChEMBL) should be curated to obtain drug–target interaction pairs; dissimilar activities should be assessed for the same drug–target pair in various studies. Databases with their own unique activity measure may be considered to have a higher curation level. Most of the databases listed offer a large panel of interactions and are not focused on a particular set of targets, genes or drugs, with the notable exception of K_i DB, which is rich in psychoactive drugs and also contains internally derived data. However, a bias toward some targets or genes might exist in all these databases, since drug affinity research in certain areas (notably cancer but also mental health thanks to efforts such as K_i DB) is more prolific.