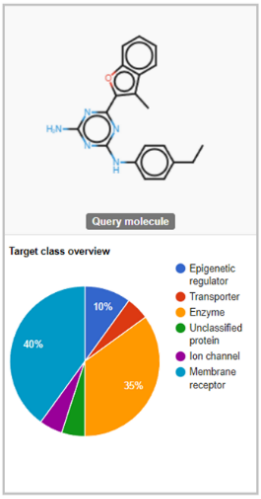
Predicting Off-targets from ChEMBL Data Using the Polypharmacology Browser

Maedeh Darsaraee and Jean-Louis Reymond

Department of Chemistry, Biochemistry and Pharmacy, University of Berne,
Freiestrasse 3, 3012 Berne, Switzerland
maedeh.darsaraee@unibe.ch

The public archive ChEMBL, which collects bioactive compounds and their associated targets from the literature, has been used by many groups to build models predicting the possible targets of small molecules to guide the experimental search for off-targets. In our group we have developed the polypharmacology browsers (PPB and PPB2),[1,2] which assign possible targets to a query molecule based on molecular fingerprint similarities to ChEMBL molecules, and provided critical insights in several practical case studies such as the identification of LPAATβ as the actual target of a putative kinase inhibitor (Figure).[3,4]



Rank	ChEMBL ID	Common name	Nearest neighbours
1	CHEMBL4772	1-acylglycerol-3-phosphate O-acyltransferase beta	Show NN
2	CHEMBL253	Cannabinoid CB2 receptor	Show NN
3	CHEMBL1904	Glutamate [NMDA] receptor subunit epsilon 2	Show NN
4	CHEMBL1800	Corticotropin releasing factor receptor 1	Show NN
5	CHEMBL224	Serotonin 2a (5-HT2a) receptor	Show NN
6	CHEMBL228	Serotonin transporter	Show NN
7	CHEMBL2916	Telomerase reverse transcriptase	Show NN
8	CHEMBL1293222	Nucleotide-binding oligomerization domain-containing protein 1	Show NN
9	CHEMBL2971	Tyrosine-protein kinase JAK2	Show NN
10	CHEMBL3764	Urotensin II receptor	Show NN
11	CHEMBL3833	Trace amine-associated receptor 1	Show NN
12	CHEMBL204	Thrombin	Show NN
13	CHEMBL3018	Matriptase	Show NN
14	CHEMBL3286	Urokinase-type plasminogen activator	Show NN
15	CHEMBL244	Coagulation factor X	Show NN

However, our PPB and PPB2 models are mostly associated only a single target per ChEMBL molecule and one protein type of targets which is a single protein. To better integrate the existing polypharmacology information available in ChEMBL, we are updating our PPB to handle multi-target information for ChEMBL molecules, using various machine learning models into account, and exploiting the latest version of the database featuring a total of 1.9 million molecule-target associations.

Keywords: Computer-aided drug design, Polypharmacology, Target prediction, Web-based tool, Cheminformatics

[1] M. Awale, J. L. Reymond, *J. Cheminf.* **2017**, *9*, 11, DOI: 10.1186/s13321-017-0199-x.
[2] M. Awale, J.-L. Reymond, *J. Chem. Inf. Model.* **2019**, *59*, 10, DOI: 10.1021/acs.jcim.8b00524.
[3] M. Poirier, M. Awale, M. A. Roelli, G. T. Giuffredi, L. Ruddigkeit, L. Evensen, A. Stooss, S. Calarco, J. B. Lorens, R.-P. Charles, J.-L. Reymond, *ChemMedChem* **2019**, *14*, 224, DOI: 10.1002/cmdc.201800554.
[4] M. R. Cunha, R. Bhardwaj, A. L. Carrel, S. Lindinger, C. Romanin, R. Parise-Filho, M. A. Hediger, J.-L. Reymond, *RSC Med. Chem.* **2020**, *11*, 1032, DOI: 10.1039/D0MD00145G.