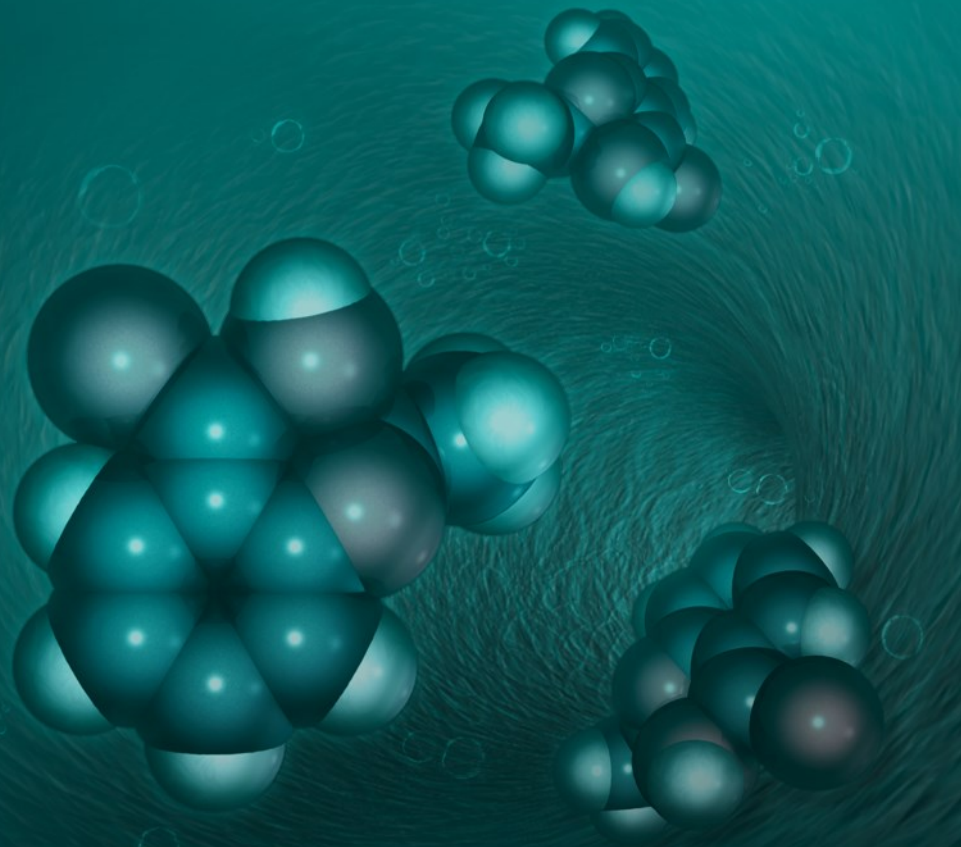
ChEMBL Database: Open data for use in Toxicity Prediction

Anne Hersey

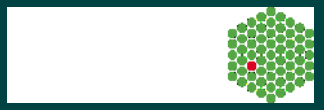
ChEMBL Group, EMBL-EBI



https://www.ebi.ac.uk/chembl

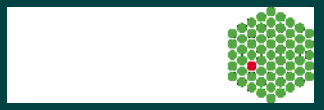
Exercises

- Target Searching – finding data on a specific protein target
- Compound Searching – finding data on substructural fragments and their associated bioactivity data
- Assay and Document searching – finding data on more complex assays from text searching



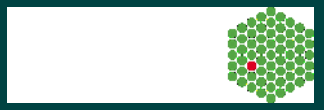
HANDS ON - Target Searching

- Exercise 1
  - You will get to know the ChEMBL Interface and the data returned from target searches
  - You will identify the experimental data available for compounds interacting with specific protein targets



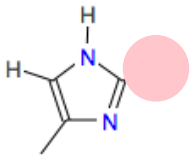
Exercise 1

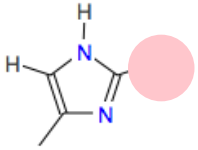
1) Search for hERG using a keyword search. How many compounds are bioactivity results are available?
2) Look at the target report card for hERG.
3) Display the Ki bioactivity data
4) Download all the Ki data into excel. Look at data (smiles, standard_values, pCHEMBL, properties, cmpds_chemblid, parent_cmpd_chemblid)
5) Other ion channels are believed to be associated with Torsade-de-Pointes arrythmias such as Nav1.5 ((SCN5A, Uniprot_id: Q14524) and Cav1.2 (CACNA1C, Uniprot_ID:Q14524). (Cardiovasc. Res., (2011),91,53-61)
6) Search for these in different ways in ChEMBL.
   Name searches (different identifiers), Browse targets, Blast Search (hint: you can get the FASTA file from Uniprot http://www.uniprot.org)



HANDS ON - Compound Searching

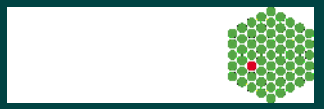
- Exercise 2
  - You will do some sub-structure searches and compare the bioactivity data for the compounds found in your searches.
  - Specifically we will investigate the difference between Cytochrome P450 inhibition of imidazoles substituted and unsubstituted between the nitrogens





Exercise 2

1. Draw the unsubstituted imidazole
2. Do a substructure search
3. Select Plot View
4. Filter using property plot for compounds with MWT<500 and ALOGP<5 (Hint: remember to update plot)
5. Select Table View
6. Display Bioactivities
7. Download to excel
8. Use excel filter to select targets (pref_name column) which are Cytochrome targets
9. Paste cytochrome set to another workbook – label all rows unsubstituted
10. Repeat 1-9 with the substituted imidazole
11. Plot data to see if there are any differences in inhibition (pCHEMBL values are a quick way of looking at this)



HANDS ON – Assay/Document Searching

Exercise 3

- You will identify potentially interesting datasets for more complex end-points by searching the assay or document descriptions
- Specifically we will search for:
- Ames test experimental data
- TG-GATES biochemistry data

Exercise 3 - Ames Search

1. Type Ames into search box and select assays
2. Look at types of data identified (you can filter out any assays that don't look relevant at this point)
3. Select to display bioactivity data
4. Use hide/unhide options to display activity_comment
5. Download data into excel

Exercise 3 - Search for TG-GATES data

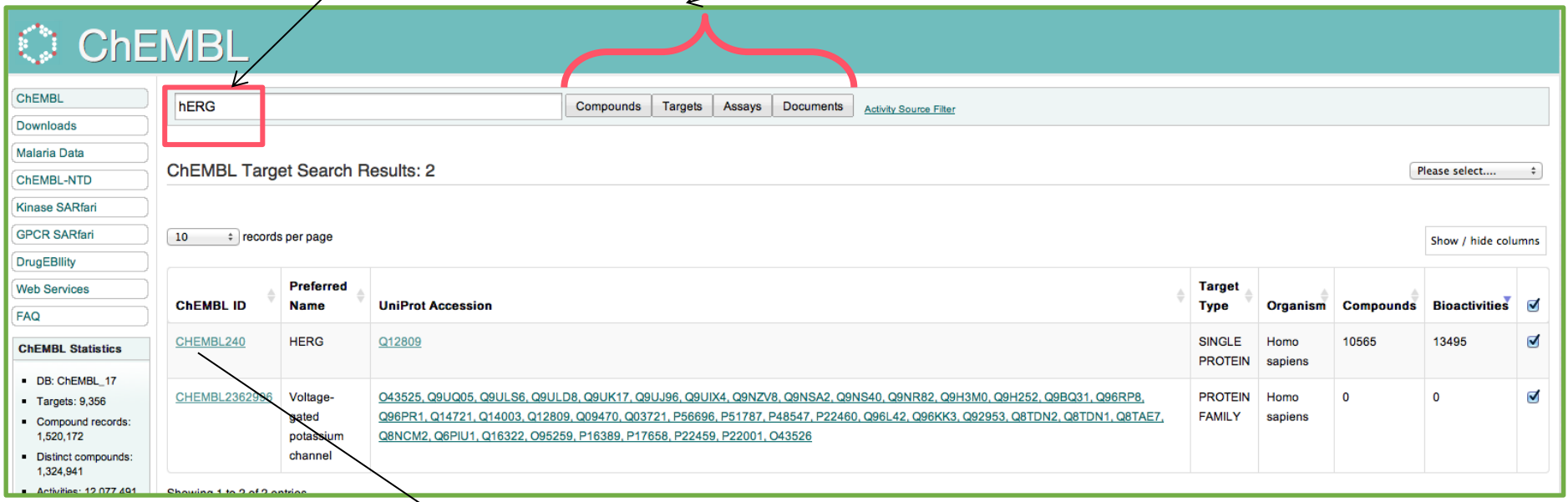
1. Select just this data source (hint: select activity source filter)
2. To search for all data for this source type * into search box and select documents
3. Select to view the biochemistry dataset via the document report card
4. Display the bioactivity data
5. Use hide/unhide to view activity comments

Hands On - Answers

Exercise 1

Input search term here

Select type of data for search



Select target to view

hERG – Target Report Card

Target Report Card

Target Name and Classification

Target ID	CHEMBL240
Target Type	SINGLE PROTEIN
Preferred Name	HERG
Synonyms	ERG | ERG-1 | ERG1 | Eag homolog | Eag-related protein 1 | Ether-a-go-go-related gene potassium channel 1 | Ether-a-go-go-related protein 1 | H-ERG | HERG | KCNH2 | Potassium voltage-gated channel subfamily H member 2 | Voltage-gated potassium channel subunit Kv11.1 | hERG-1 | hERG1
Organism	Homo sapiens
Species Group	No
Protein Target Classification	ion channel|vgc|vgc|volt|cationic|k|kcnh, kv10-12.x (ether-a-go-go)

Target Components

Component Description	Relationship	Accession
Potassium voltage-gated channel subfamily H member 2	SINGLE PROTEIN	Q12809

Target Relations

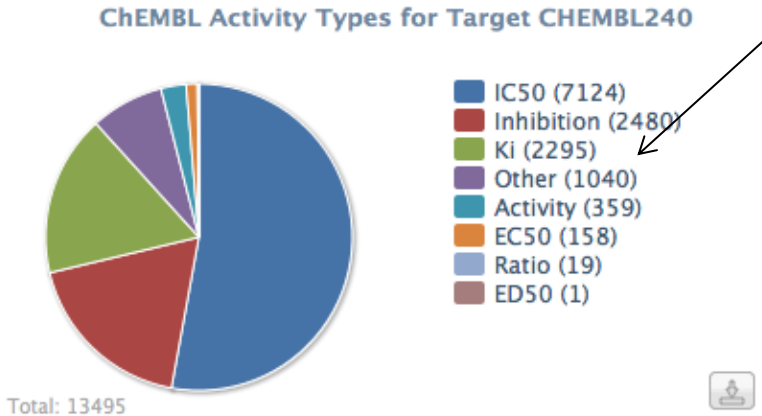
ChEMBL ID	Pref Name	Target Type
CHEMBL2362996	Voltage-gated potassium channel	PROTEIN FAMILY

Approved Drugs

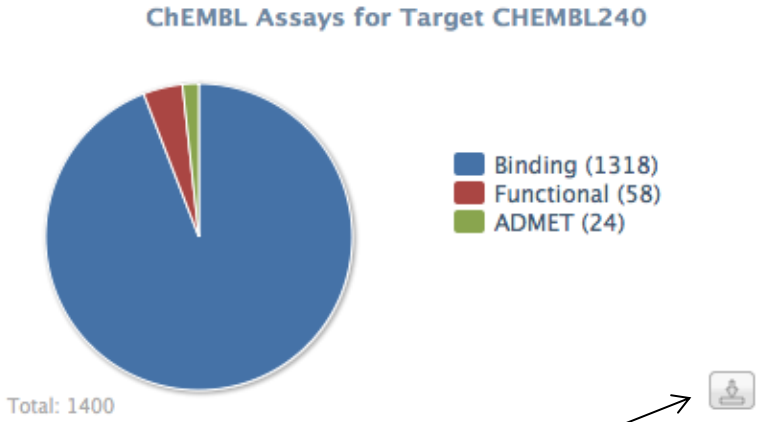
ChEMBL ID	Name	Mechanism of Action	References
CHEMBL1083993	AMIODARONE HYDROCHLORIDE	HERG blocker	DailyMed
CHEMBL473	DOFETILIDE	HERG blocker	DailyMed
CHEMBL1201729	DRONEDARONE HYDROCHLORIDE	HERG blocker	Expert FDA
CHEMBL1200564	IBUTILIDE FUMARATE	HERG blocker	DailyMed PubMed
CHEMBL1700	SOTALOL HYDROCHLORIDE	HERG blocker	DailyMed

Target Associated Bioactivities

Click here to see Ki data



Target Associated Assays



Target Ligand Efficiencies

Click here to see all data

Exercise 1- bioactivity data

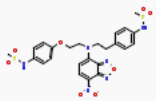
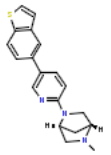
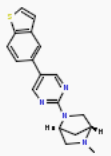
ChEMBL Bioactivity Search Results: 2295

Please select....
Download All Data (TAB)
Download All Data (XLS)

Please select....

100 records per page

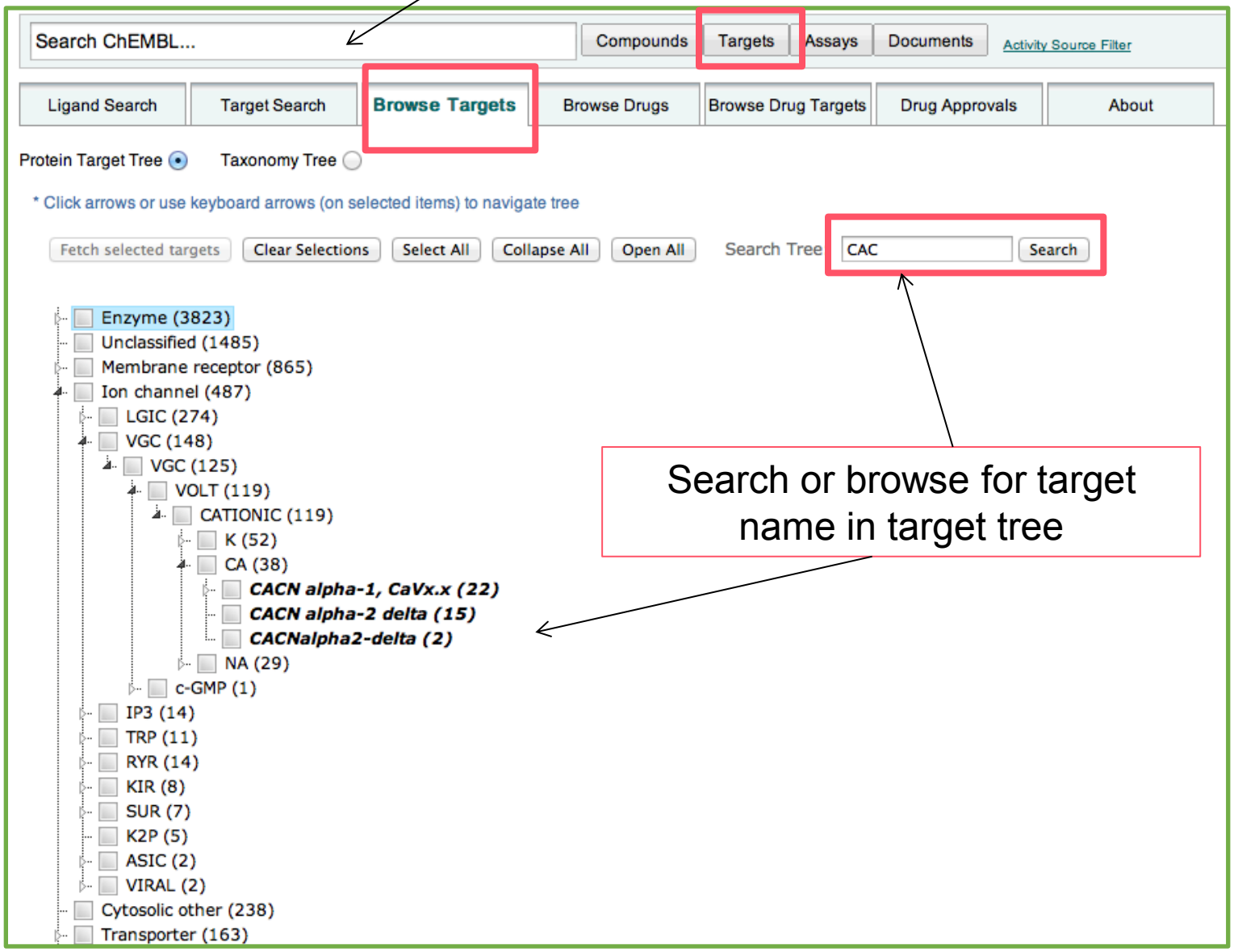
Show / hide columns

Ingredient	Molweight	Standard Type	Relation	Standard Value	Standard Units	pChEMBL Value	Assay Type	Description	Assay Src Description	Assay Organism	Target Type	Target Name	Target Organism	Reference
 CHEMBL227134	590.63	Ki	=	.39	nM	9.41	B	Displacement of [3H]dofetilide from hERG expressed in HEK293 cells by SPA	Scientific Literature	Homo sapiens	SINGLE PROTEIN	HERG	Homo sapiens	J. Med. Chem., (2007) 50:13:2931
 CHEMBL1085090	321.44	Ki	=	.55	nM	9.26	B	Displacement of [3H]dofetidile from human ERG by whole-cell patch clamp	Scientific Literature	Homo sapiens	SINGLE PROTEIN	HERG	Homo sapiens	Bioorg. Med. Chem. Lett., (2010) 20:12:3636
 CHEMBL1085806	322.43	Ki	=	.56	nM	9.25	B	Displacement of [3H]dofetidile from human ERG by whole-cell patch clamp	Scientific Literature	Homo sapiens	SINGLE PROTEIN	HERG	Homo sapiens	Bioorg. Med. Chem. Lett., (2010) 20:12:3636

Download file contains:
Parent smiles, parent and salt identifiers
Physchem descriptors
Ki values
pCHEMBL values - -log10(molar IC50, XC50, EC50, AC50, Ki, Kd or Potency)

Exercise 1 – searching for other ion channels

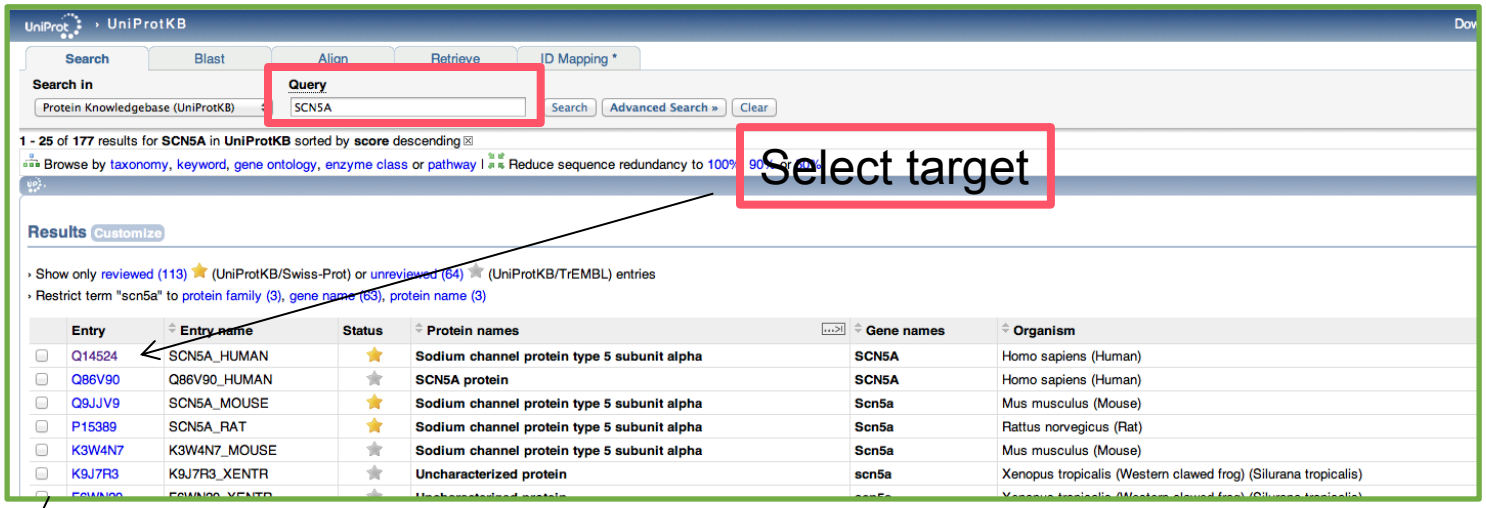
Name search for target name, CHEMBL_ID or Uniprot_ID

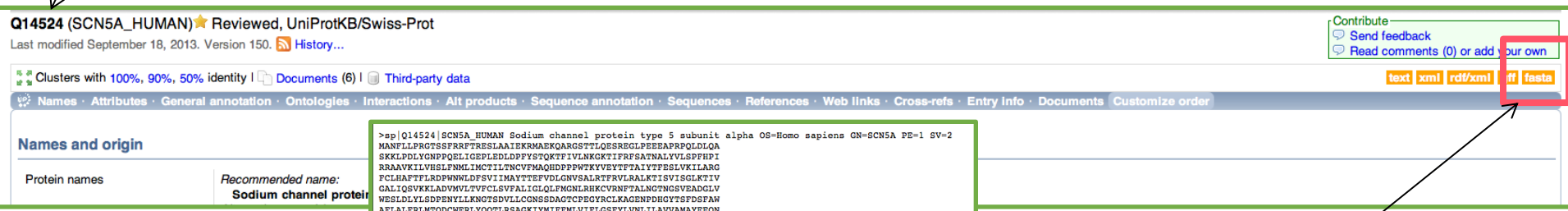


Search or browse for target name in target tree

Exercise 1 – BLAST search

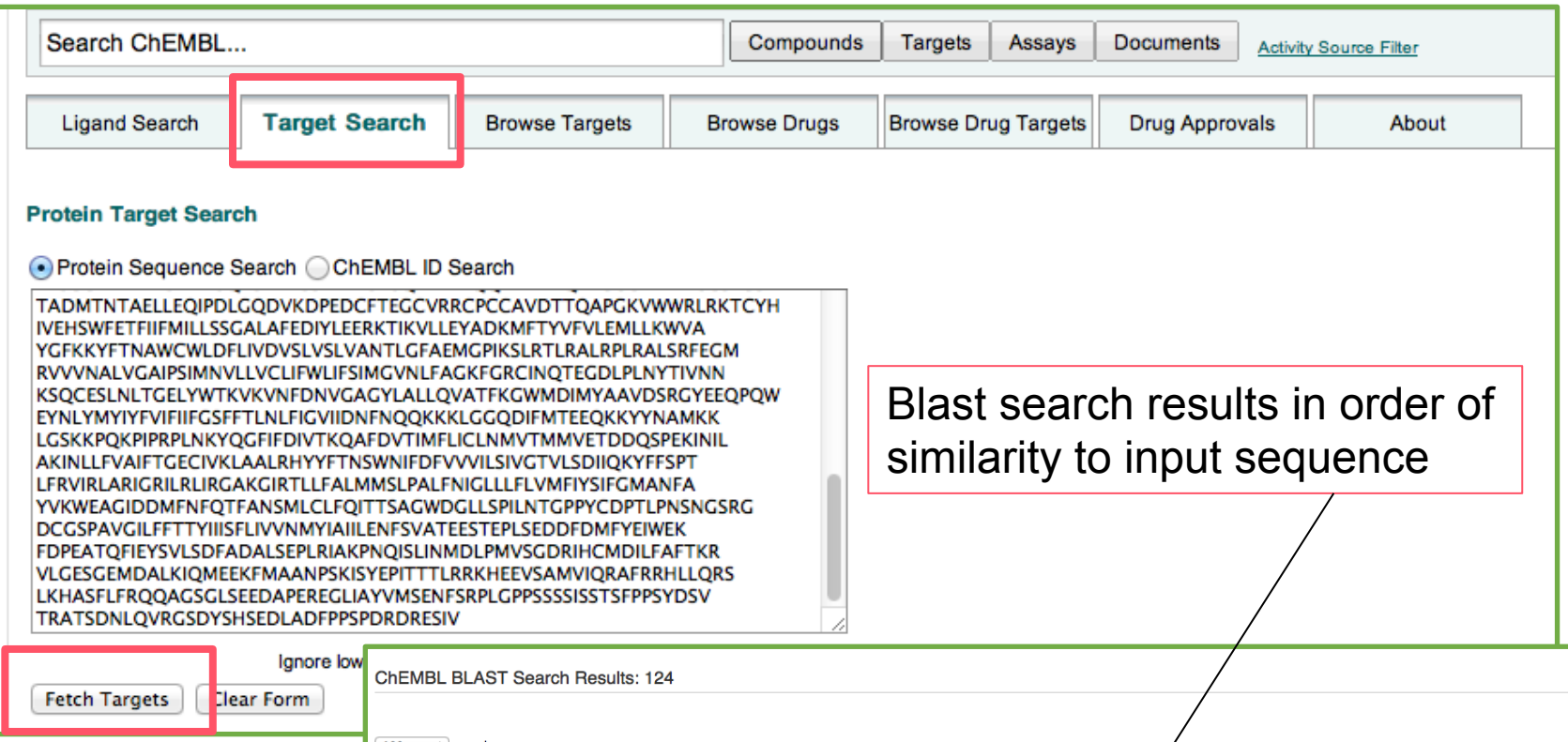
Find target in Uniprot http://www.uniprot.org



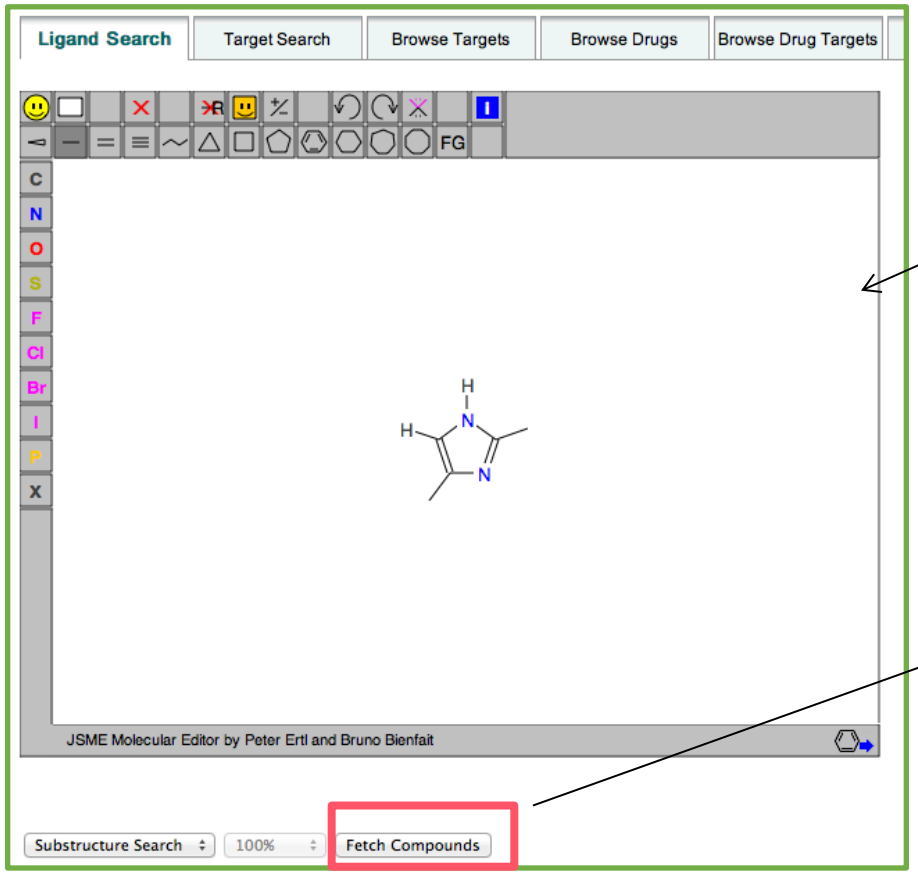



```
>sp|Q14524|SCN5A_HUMAN Sodium channel protein type 5 subunit alpha OS=Homo sapiens GN=SCN5A PE=1 SV=2
MANFLLPRGTSSFRRFTRESLAAIEKRMAEKQARGSTTLQESREGLPEEEAPRPQLDLQA
SKKLPDLYGNPPQELIGEPLEDLDPFYSTQKTFIVLNKGKTIFRFSATNALYVLSPFHPI
RRAAVKILVHSLFNMLIMCTILTNCVFMAQHDPPPWTKYVEYTFTAIYTFESLVKILARG
FCLHAFTFLRDPWNWLDFSVIIMAYTTEFVDLGNVSALRTFRVLRALKTISVISGLKTIV
GALIQSVKKLADVMVLTVFCLSVFALIGLQLFMGNLRHKCVRNFTALNGTNGSVEADGLV
WESLDLYLSDPENYLLKNGTSDVLLCGNSSDAGTCPEGYRCLKAGENPDHGYTSFDSFAW
AFLALFRLMTQDCWERLYQQTLRSAGKIYMIFFMLVIFLGSFYLVNLILAVVAMAYEEQN
QATIAETEEKEKRFQEAMEMLKKEHEALTIRGVDTVSRSSLEMSPLAPVNSHERRSKRRK
RMSSGTEECGEDRLPKSDSEDGPRAMNHLSLTRGLSRTSMKPRSSRGSIFTFRRRDLGSE
ADFADDENSTAGESESHHTSLLVPWPLRRTSAQGQPSPGTSAPGHALHGKKNSTVDCNGV
VSLLGAGDPEATSPGSHLLRPVMLEHPPDTTTPSEEPGGPQMLTSQAPCVDGFEEPGARQ
RALSAVSVLTSALEELEESRHKCPPCWNRLAQRYLIWECCPLWMSIKQGVKLVVMDPFTD
LTITMCIVLNTLFMALEHYNMTSEFEEMLQVGNLVFTGIFTAEMTFKIIALDPYYYFQQG
WNIFDSIIVILSLMELGLSRMSNLSVLRSFRLLRVFKLAKSWPTLNTLIKIIGNSVGALG
NLTLVLAIIVFIFAVVGMQLFGKNYSELRDSDSGLLPRWHMMDFFHAFLIIFRILCGEWI
ETMWDCMEVSGQSLCLLVFLLVMVIGNLVVLNLFLALLLSSFSADNLTAPDEDREMNNLQ
LALARIQRGLRFVKRTTWDFCCGLLRQRPQKPAALAAQGQLPSCIATPYSPPPPETEKVP
PTRKETFEEGEQPGQGTPGDPEPVCVPIAVAESDTDDQEEDEENSLGTEEESSKQQESQ
PVSGGPEAPPDSRTWSQVSATASSEAEASASQADWRQQWKAEPQAPGCGETPEDSCSEGS
TADMTNTAELLEQIPDLGQDVKDPEDCFTEGCVRRCPCCAVDTTQAPGKVWWRLRKTCYH
IVEHSWFETFIIFMILLSSGALAFEDIYLEERKTIKVLLEYADKMFTYVFVLEMLLKWVA
YGFKKYFTNAWCWLDFLIVDVSLVSLVANTLGFAEMGPIKSLRTLRALRPLRALSRFEGM
RVVVNALVGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFGRCINQTEGDLPLNYTIVNN
KSQCESLNLTGELYWTKVKVNFDNVGAGYLALLQVATFKGWMDIMYAAVDSRGYEEQPQW
EYNLYMYIYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKLGGQDIFMTEEQKKYYNAMKK
LGSKKPQKPIPRPLNKYQGFIFDIVTKQAFDVTIMFLICLNMVTMMVETDDQSPEKINIL
AKINLLFVAIFTGECIVKLAALRHYYFTNSWNIFDFVVVILSIVGTVLSDIIQKYFFSPT
LFRVIRLARIGRILRLIRGAKGIRTLLFALMMSLPALFNIGLLLFLVMFIYSIFGMANFA
YVKWEAGIDDMFNFQTFANSMLCLFQITTSAGWDGLLSPILNTGPPYCDPTLPNSNGSRG
DCGSPAVGILFFTTYIIISFLIVVNMYIAILLENFSVATEESTEPLSEDDFDMFYEIWEK
FDPEATQFIEYSVLSDFADALSEPLRIAKPNQISLINMDLPMVSGDRIHCMDILFAFTKR
VLGESGEMDALKIQMEEKFMAANPSKISYEPITTTLRRKHEEVSAMVIQRAFRRHLLQRS
LKHASFLFRQQAGSGLSEEDAPEREGLIAYVMSENFSRPLGPPSSSSISSTSFPPSYDSV
TRATSDNLQVRGSDYSHSEDLADFPPSPDRDRESIV
```

Select sequence from FASTA file


Search ChEMBL...
Compounds
Targets
Assays
Documents
Activity Source Filter
Ligand Search
Target Search
Browse Targets
Browse Drugs
Browse Drug Targets
Drug Approvals
About
Protein Target Search
Protein Sequence Search
ChEMBL ID Search
TADMTNTAELLEQIPDLGQDVKDPEDCFTEGCVRRCPCCAVDTTQAPGKVWWRLRKTCYH
IVEHSWFETFIIFMILLSSGALAFEDIYLEERKTIKVLLEYADKMFTYVFVLEMLLKWVA
YGFKKYFTNAWCWLDFLIVDVSLVSLVANTLGFAEMGPIKSLRTLRALRPLRALSRFEGM
RVVVNALVGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFGRCINQTEGDLPLNYTIVNN
KSQCESLNLTGELYWTKVKVNFDNVGAGYLALLQVATFKGWMDIMYAAVDSRGYEEQPQW
EYNLYMYIYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKLGGQDIFMTEEQKKYYNAMKK
LGSKKPQKPIPRPLNKYQGFIFDIVTKQAFDVTIMFLICLNMVTMMVETDDQSPEKINIL
AKINLLFVAIFTGECIVKLAALRHYYFTNSWNIFDFVVVILSIVGTVLSDIIQKYFFSPT
LFRVIRLARIGRILRLIRGAKGIRTLLFALMMSLPALFNIGLLLFLVMFIYSIFGMANFA
YVKWEAGIDDMFNFQTFANSMLCLFQITTSAGWDGLLSPILNTGPPYCDPTLPNSNGSRG
DCGSPAVGILFFTTYIIISFLIVVNMYIAIILENFSVATEESTEPLSEDDFDMFYEIWEK
FDPEATQFIEYSVLSDFADALSEPLRIAKPNQISLINMDLPMVSGDRIHCMDILFAFTKR
VLGESGEMDALKIQMEEKFMAANPSKISYEPITTTLRRKHEEVSAMVIQRAFRRHLLQRS
LKHASFLFRQQAGSGLSEEDAPEREGLIAYVMSENFSRPLGPPSSSSISSTSFPPSYDSV
TRATSDNLQVRGSDYSHSEDLADFPPSPDRDRESIV
Ignore low
Fetch Targets
Clear Form
Blast search results in order of similarity to input sequence
ChEMBL BLAST Search Results: 124
Please select....
100
records per page
Show / hide columns
Query Sequence
ChEMBL ID
Preferred Name
UniProt Accession
Target Type
Organism
Compounds
Endpoints
% Identity
BLAST Score
E-Value
sp|Q14524|SCN5A_HUMAN
CHEMBL1980
Sodium channel protein type V alpha subunit
Sodium channel protein type V alpha subunit
SINGLE PROTEIN
Homo sapiens
493
588
100
4177
0e+0
sp|Q14524|SCN5A_HUMAN
CHEMBL2331043
Sodium channel alpha subunit
Sodium channel alpha subunit
PROTEIN FAMILY
Homo sapiens
0
0
100
4177
0e+0
sp|Q14524|SCN5A_HUMAN
CHEMBL1770040
Voltage-gated sodium channel type V alpha polypeptide
Voltage-gated sodium channel type V alpha polypeptide
SINGLE PROTEIN
Oryctolagus cuniculus
1
1
96.19
584
0e+0
sp|Q14524|SCN5A_HUMAN
CHEMBL2362984
Voltage-sensitive sodium channel alpha-subunit
Voltage-sensitive sodium channel alpha-subunit
SINGLE PROTEIN
Pediculus humanus
0
0
55.35
842
0e+0
sp|Q14524|SCN5A_HUMAN
CHEMBL3866
Sodium channel protein type V alpha subunit
Sodium channel protein type V alpha subunit
SINGLE PROTEIN
Rattus norvegicus
69
71
94.01
3792
0e+0
sp|Q14524|SCN5A_HUMAN
CHEMBL2096682
Sodium channel alpha subunits; brain (Types I, II, III)
Sodium channel alpha subunits; brain (Types I, II, III)
PROTEIN FAMILY
Homo sapiens
212
374
62.28
2433
0e+0
sp|Q14524|SCN5A_HUMAN
CHEMBL4187
Sodium channel protein type II alpha subunit
Sodium channel protein type II alpha subunit
SINGLE PROTEIN
Homo sapiens
142
244
62.28
2433
0e+0
sp|Q14524|SCN5A_HUMAN
CHEMBL2095171
Sodium channel alpha subunits; brain (Types I, II, III)
Sodium channel alpha subunits; brain (Types I, II, III)
PROTEIN FAMILY
Rattus norvegicus
204
344
62.07
2427
0e+0

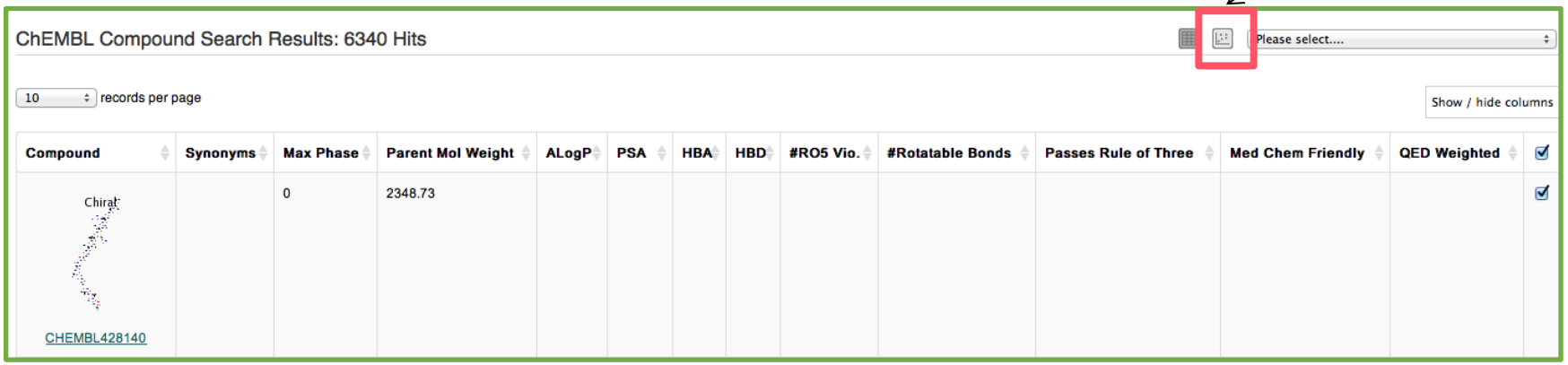
Exercise 2 – Substructure search



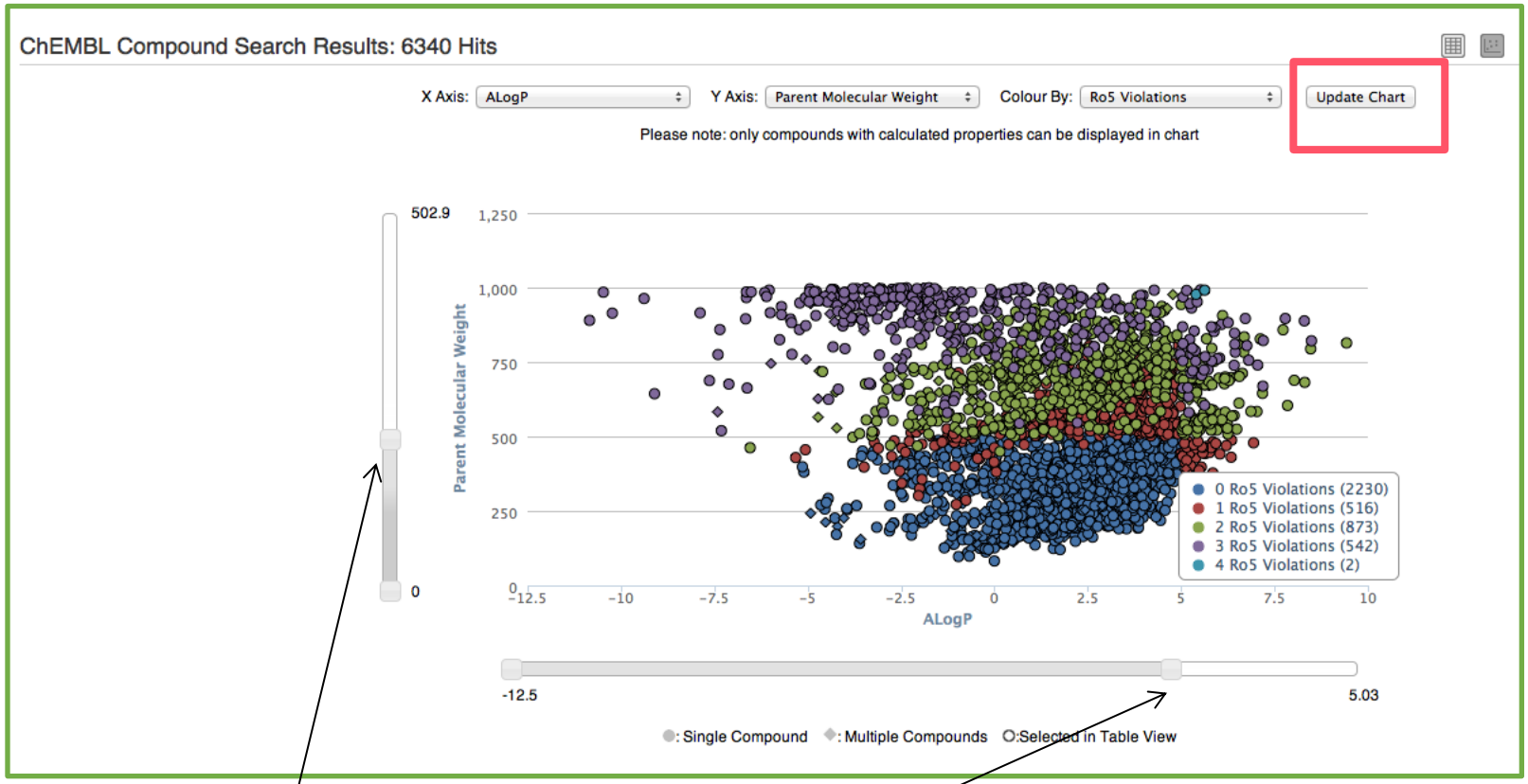
Draw substructure

Search returns 6340 compounds

Select chart view



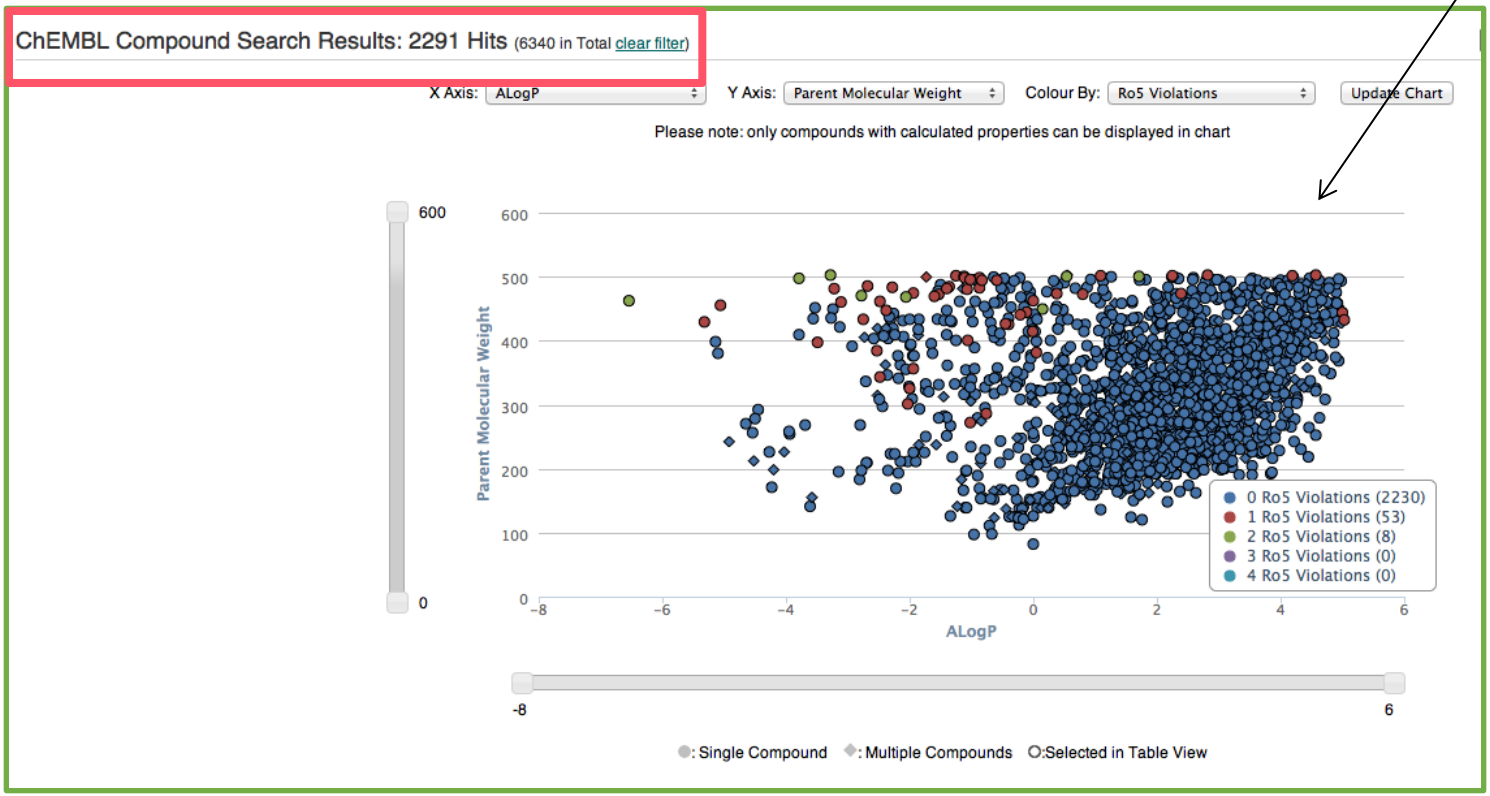
Exercise 2 – Filter on compound properties



Use sliders to filter on MWT and LogP

Exercise 2 – Filtered dataset

Displaying only compounds with MWT<500 and LogP<5



Exercise 2 – Table View of Filtered Compound Set

Exercise 2 – Bioactivity data on filtered compounds

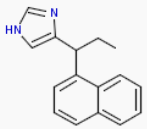
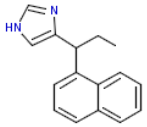
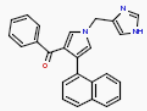
Download data into excel

ChEMBL Bioactivity Search Results: 14028

Please select....

Please select....
Download All Data (TAB)
Download All Data (XLS)

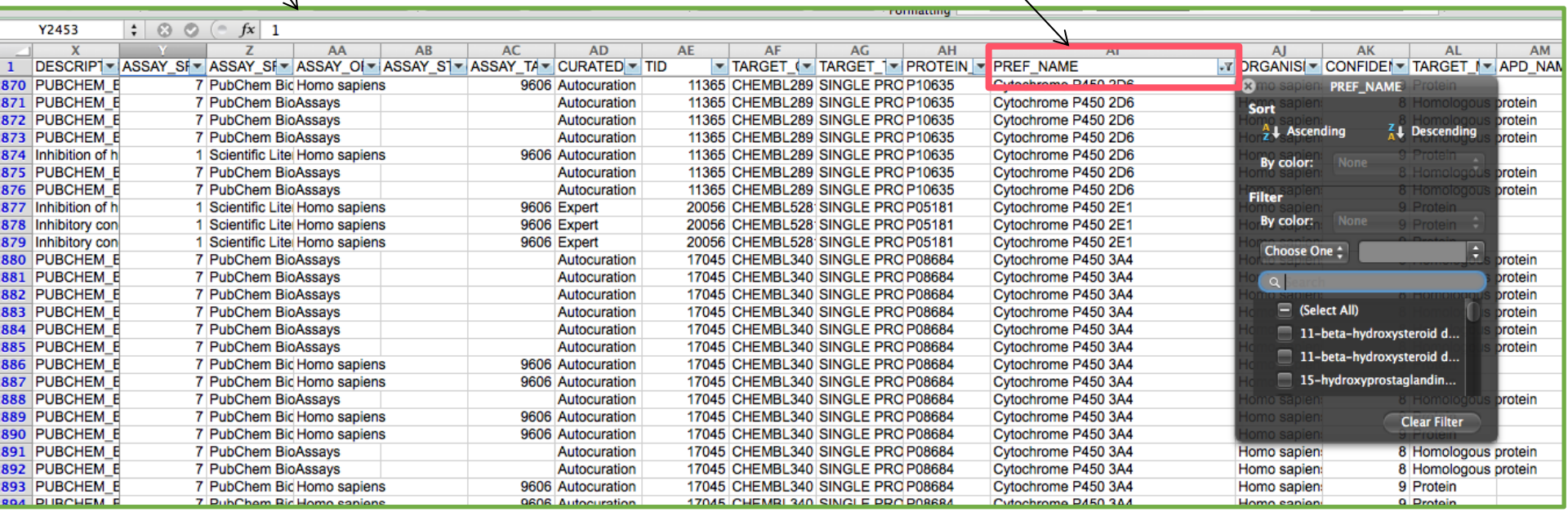
10 records per page

Ingredient	Molweight	Standard Type	Relation	Standard Value	Standard Units	Assay Type	Description	Assay Src Description	Assay Organism	Target Type	Target Name	Organism	Reference
 CHEMBL102076	236.31	Ratio	=	3.48		B	Selectivity ratio of Ki of alpha-1 adrenergic receptor to alpha-2 receptor	Scientific Literature		UNCHECKED	Unchecked		J. Med. Chem., (1996) 39:15:3001
 CHEMBL102076	236.31	Ki	=	165	nM	B	Alpha-2 adrenergic receptor binding affinity was tested against membrane preparations of rat brain using 0.2 nM [3H]rauwolscine as radioligand	Scientific Literature		PROTEIN FAMILY	Adrenergic receptor alpha-2	Rattus norvegicus	J. Med. Chem., (1996) 39:15:3001
 CHEMBL102187	377.44	IC50	=	115	nM	F	Inhibition of [3H]FPP incorporation into Ha-ras protein by Farnesyltransferase	Scientific Literature		PROTEIN COMPLEX	Protein farnesyltransferase	Homo sapiens	Bioorg. Med. Chem. Lett., (2001) 11:22:2963

Exercise 2 – Select just Cytochrome P450 data in excel

Excel table of results

Set filter on pref_name
Data - filter



	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	PREF_NAME	AJ	AK	AL	AM
1	DESCRIPT	ASSAY_SF	ASSAY_SF	ASSAY_OI	ASSAY_ST	ASSAY_TA	CURATED	TID	TARGET_(	TARGET_T	PROTEIN_	PREF_NAME	ORGANISI	CONFIDEI	TARGET_I	APD_NAM
870	PUBCHEM_E	7	PubChem Bic	Homo sapiens		9606	Autocuration	11365	CHEMBL289	SINGLE PRO	P10635	Cytochrome P450 2D6	Homo sapien:	8	Homologous protein	
871	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	11365	CHEMBL289	SINGLE PRO	P10635	Cytochrome P450 2D6	Homo sapien:	8	Homologous protein	
872	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	11365	CHEMBL289	SINGLE PRO	P10635	Cytochrome P450 2D6	Homo sapien:	8	Homologous protein	
873	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	11365	CHEMBL289	SINGLE PRO	P10635	Cytochrome P450 2D6	Homo sapien:	8	Homologous protein	
874	Inhibition of h	1	Scientific Lite	Homo sapiens		9606	Autocuration	11365	CHEMBL289	SINGLE PRO	P10635	Cytochrome P450 2D6	Homo sapien:	9	Protein	
875	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	11365	CHEMBL289	SINGLE PRO	P10635	Cytochrome P450 2D6	Homo sapien:	8	Homologous protein	
876	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	11365	CHEMBL289	SINGLE PRO	P10635	Cytochrome P450 2D6	Homo sapien:	8	Homologous protein	
877	Inhibition of h	1	Scientific Lite	Homo sapiens		9606	Expert	20056	CHEMBL528	SINGLE PRO	P05181	Cytochrome P450 2E1	Homo sapien:	9	Protein	
878	Inhibitory con	1	Scientific Lite	Homo sapiens		9606	Expert	20056	CHEMBL528	SINGLE PRO	P05181	Cytochrome P450 2E1	Homo sapien:	9	Protein	
879	Inhibitory con	1	Scientific Lite	Homo sapiens		9606	Expert	20056	CHEMBL528	SINGLE PRO	P05181	Cytochrome P450 2E1	Homo sapien:	9	Protein	
880	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
881	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
882	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
883	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
884	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
885	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
886	PUBCHEM_E	7	PubChem Bic	Homo sapiens		9606	Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	9	Protein	
887	PUBCHEM_E	7	PubChem Bic	Homo sapiens		9606	Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	9	Protein	
888	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
889	PUBCHEM_E	7	PubChem Bic	Homo sapiens		9606	Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	9	Protein	
890	PUBCHEM_E	7	PubChem Bic	Homo sapiens		9606	Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	9	Protein	
891	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
892	PUBCHEM_E	7	PubChem Bio	Assays			Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	8	Homologous protein	
893	PUBCHEM_E	7	PubChem Bic	Homo sapiens		9606	Autocuration	17045	CHEMBL340	SINGLE PRO	P08684	Cytochrome P450 3A4	Homo sapien:	9	Protein	

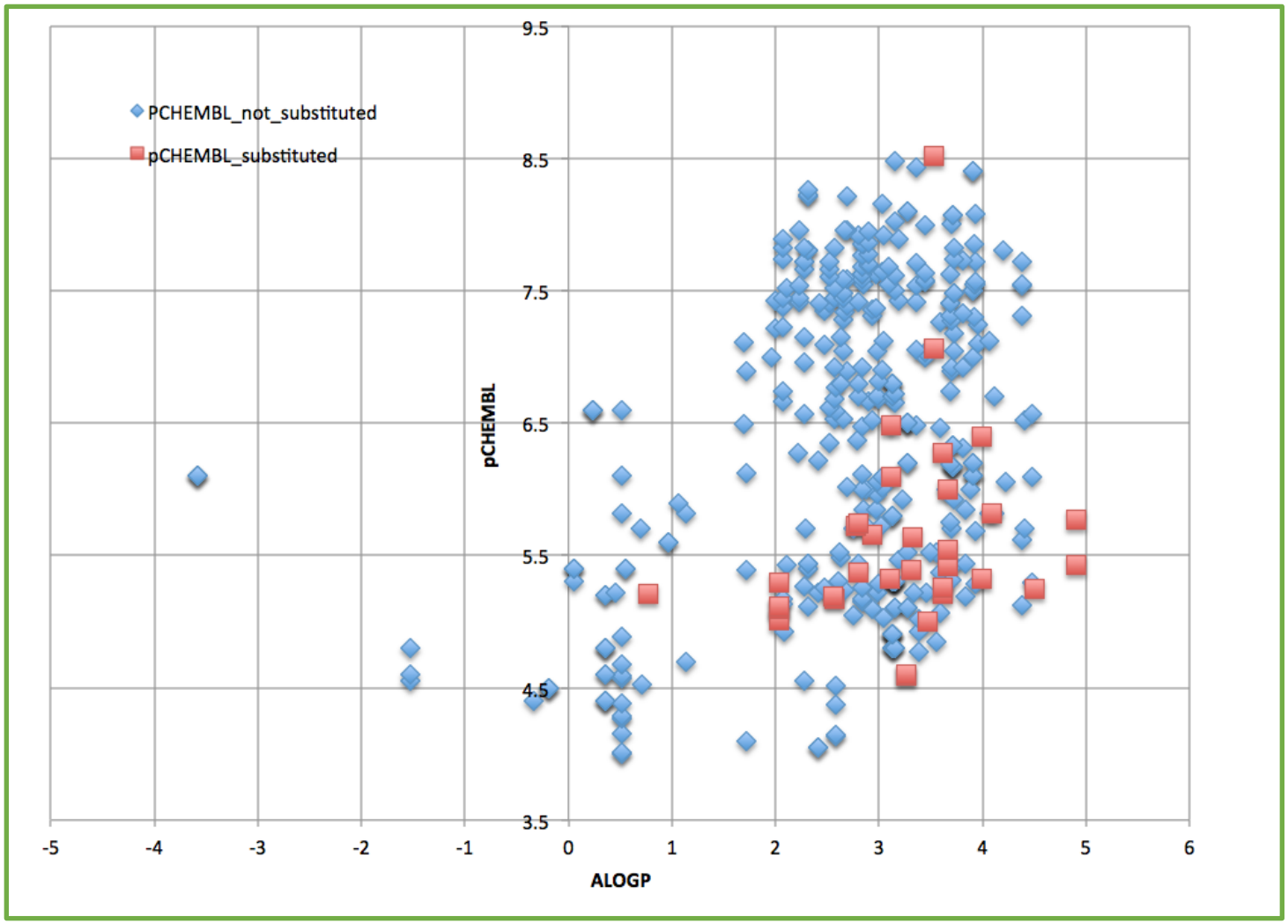
After filtering in excel 548 values

Exercise 2 - Results from substituted imidazole search

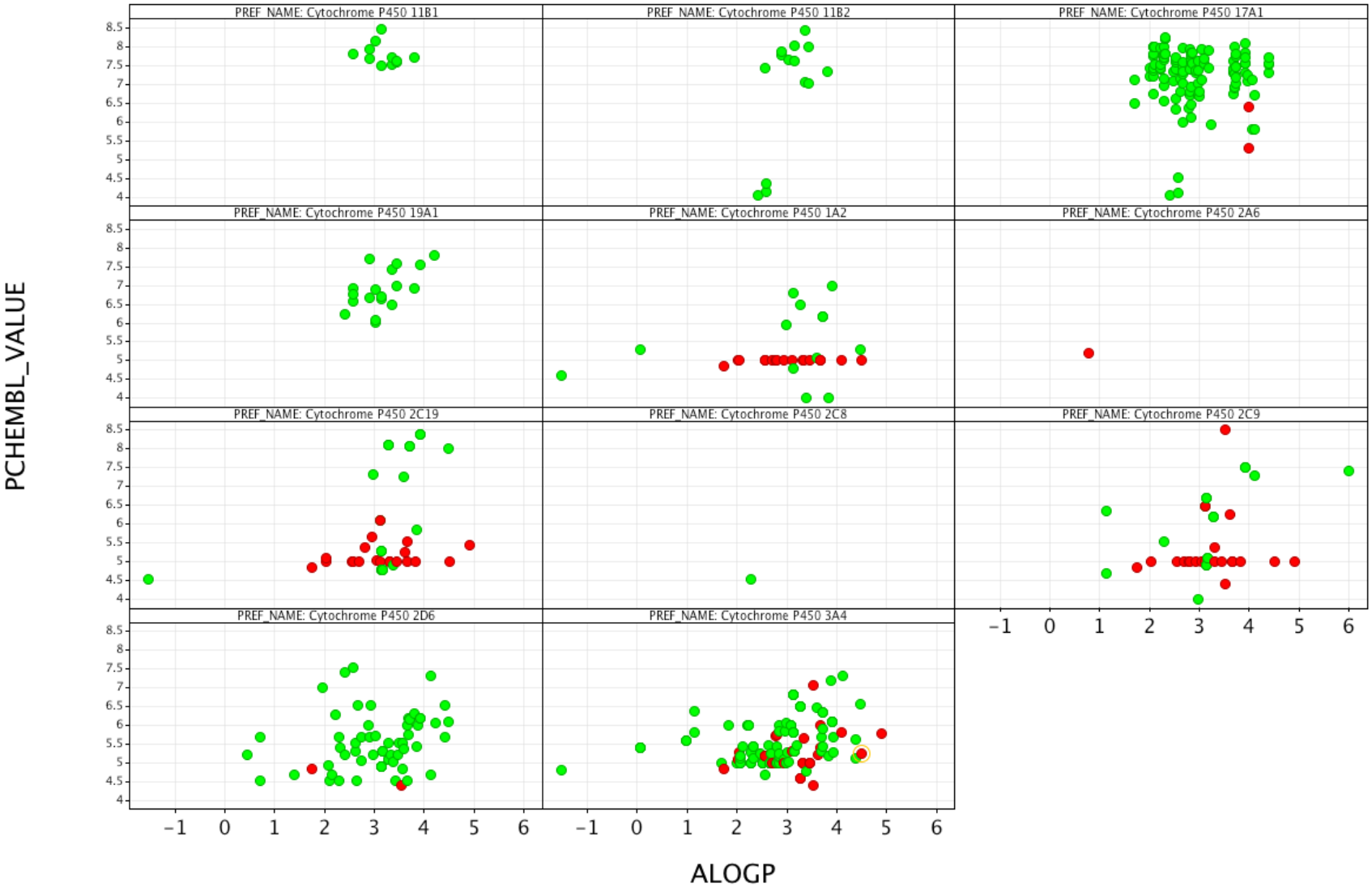
- 1297 compounds
- 1035 compounds after ALOGP <5 and MWT <500 filter
- 8028 bioactivity values
- Filter only on Cytochromes 132 bioactivity values

File: imidazole_example_final.xlsx

Exercise 2 – View of Cytochrome P450 data on substituted and unsubstituted Imidazoles

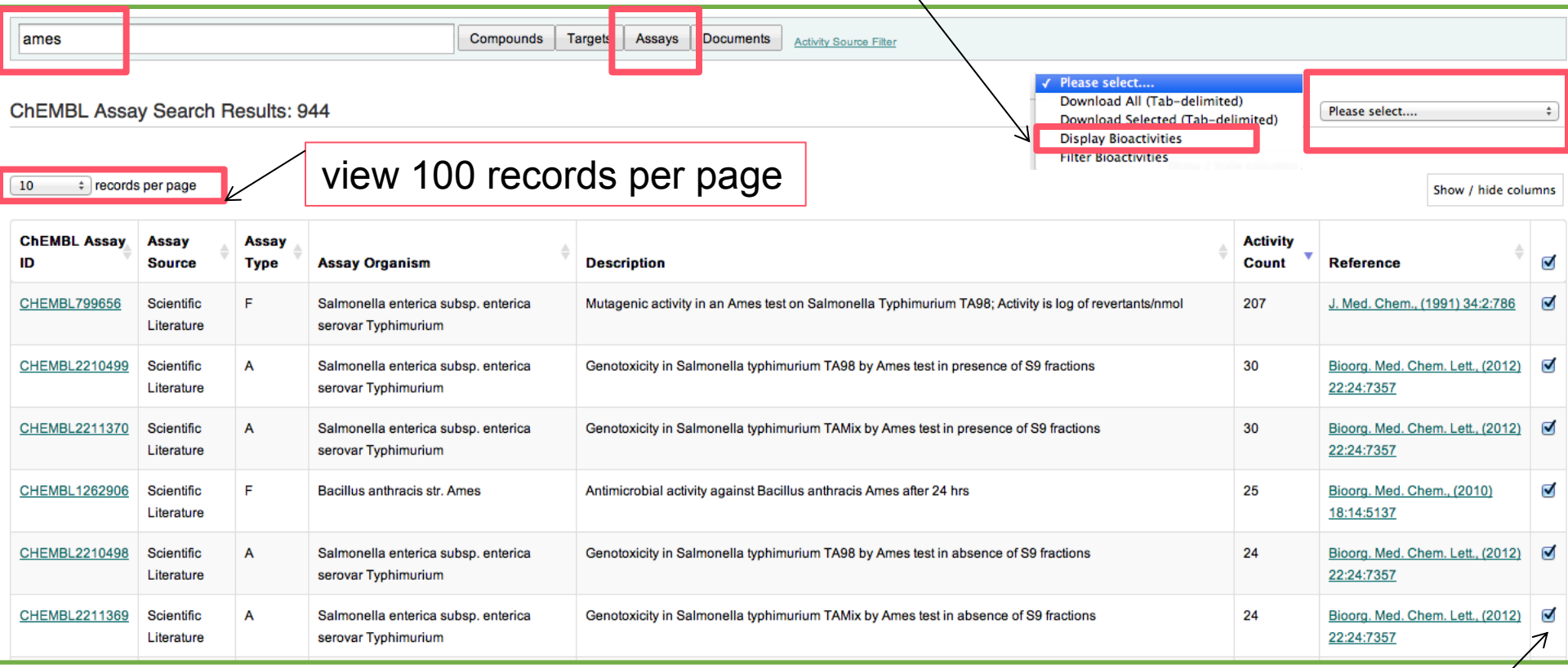


Note also includes assignment of inactive (pCHEMBL=5) for appropriate compounds



Exercise 3 – Assay search for Ames

Select “display bioactivities” to see data



ChEMBL Assay ID	Assay Source	Assay Type	Assay Organism	Description	Activity Count	Reference
CHEMBL799656	Scientific Literature	F	Salmonella enterica subsp. enterica serovar Typhimurium	Mutagenic activity in an Ames test on Salmonella Typhimurium TA98; Activity is log of revertants/nmol	207	J. Med. Chem., (1991) 34:2:786
CHEMBL2210499	Scientific Literature	A	Salmonella enterica subsp. enterica serovar Typhimurium	Genotoxicity in Salmonella typhimurium TA98 by Ames test in presence of S9 fractions	30	Bioorg. Med. Chem. Lett., (2012) 22:24:7357
CHEMBL2211370	Scientific Literature	A	Salmonella enterica subsp. enterica serovar Typhimurium	Genotoxicity in Salmonella typhimurium TAMix by Ames test in presence of S9 fractions	30	Bioorg. Med. Chem. Lett., (2012) 22:24:7357
CHEMBL1262906	Scientific Literature	F	Bacillus anthracis str. Ames	Antimicrobial activity against Bacillus anthracis Ames after 24 hrs	25	Bioorg. Med. Chem., (2010) 18:14:5137
CHEMBL2210498	Scientific Literature	A	Salmonella enterica subsp. enterica serovar Typhimurium	Genotoxicity in Salmonella typhimurium TA98 by Ames test in absence of S9 fractions	24	Bioorg. Med. Chem. Lett., (2012) 22:24:7357
CHEMBL2211369	Scientific Literature	A	Salmonella enterica subsp. enterica serovar Typhimurium	Genotoxicity in Salmonella typhimurium TAMix by Ames test in absence of S9 fractions	24	Bioorg. Med. Chem. Lett., (2012) 22:24:7357

Deselect any data that looks irrelevant

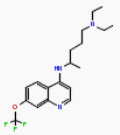
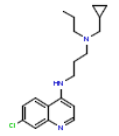
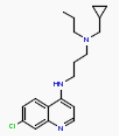
Exercise 3 – bioactivity results

Select fields to view

ChEMBL Bioactivity Search Results: 2552

Please select....

10 records per page

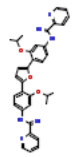
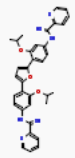
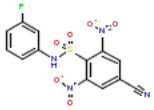
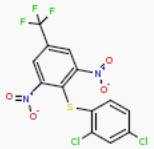
Ingredient	Molweight	Standard Type	Relation	Standard Value	Standard Units	Assay Type	Description	Assay Src Description	Assay Organism	Target Type	Target Name	Target Organism	R
 CHEMBL1087464	369.42	Activity				A	Mutagenic activity in Salmonella Typhimurium TA98 assessed as number of revertants after 48 hrs by Ames test	Scientific Literature	Salmonella enterica subsp. enterica serovar Typhimurium	ADMET	ADMET		J. 5
 CHEMBL1087469	331.88	Activity				A	Mutagenic activity in Salmonella Typhimurium TA98 assessed as number of revertants after 48 hrs by Ames test	Scientific Literature	Salmonella enterica subsp. enterica serovar Typhimurium	ADMET	ADMET		J. 5
 CHEMBL1087469	331.88	Activity				A	Mutagenic activity in Salmonella Typhimurium TA100 assessed as number of revertants after 48 hrs by Ames test	Scientific Literature	Salmonella enterica subsp. enterica serovar Typhimurium	ADMET	ADMET		J. 5

Show / hide columns

- [x] Ingredient
- [ ] Molregno
- [ ] Parent
- [ ] Parent Molregno
- [ ] Molecule Name
- [ ] Name in Reference
- [x] Molweight
- [ ] AlogP
- [ ] PSA
- [ ] #RO5 Vio.
- [ ] Canonical SMILES
- [ ] Activity Id
- [x] Standard Type
- [x] Relation
- [x] Standard Value
- [x] Standard Units
- [ ] pChEMBL Value
- [x] Activity Comment
- [ ] Data Validity
- [ ] Potential Duplicate

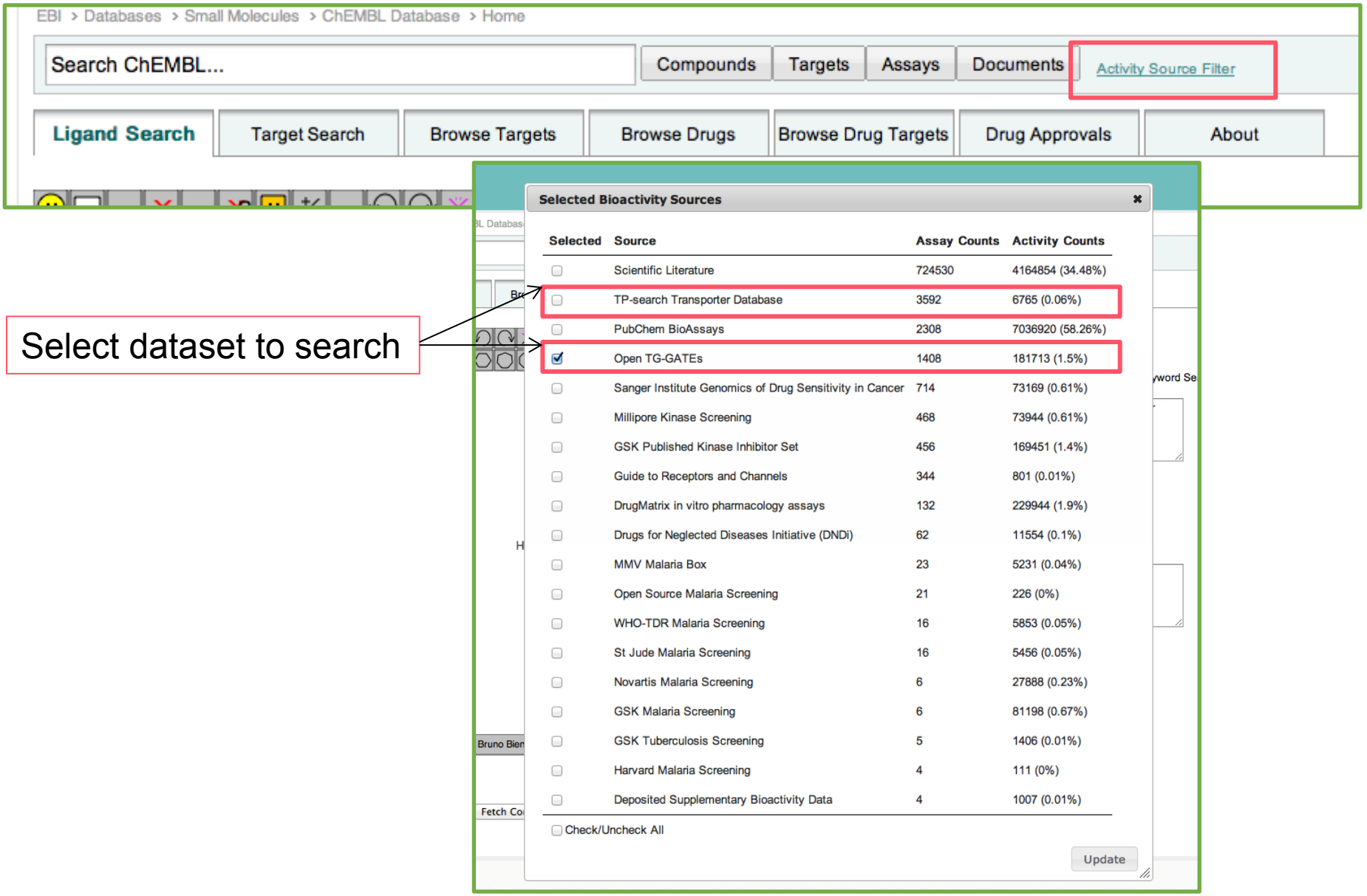
For complex data activity comment often contains qualitative information e.g active/inactive

Exercise 3 – Bioactivity Results

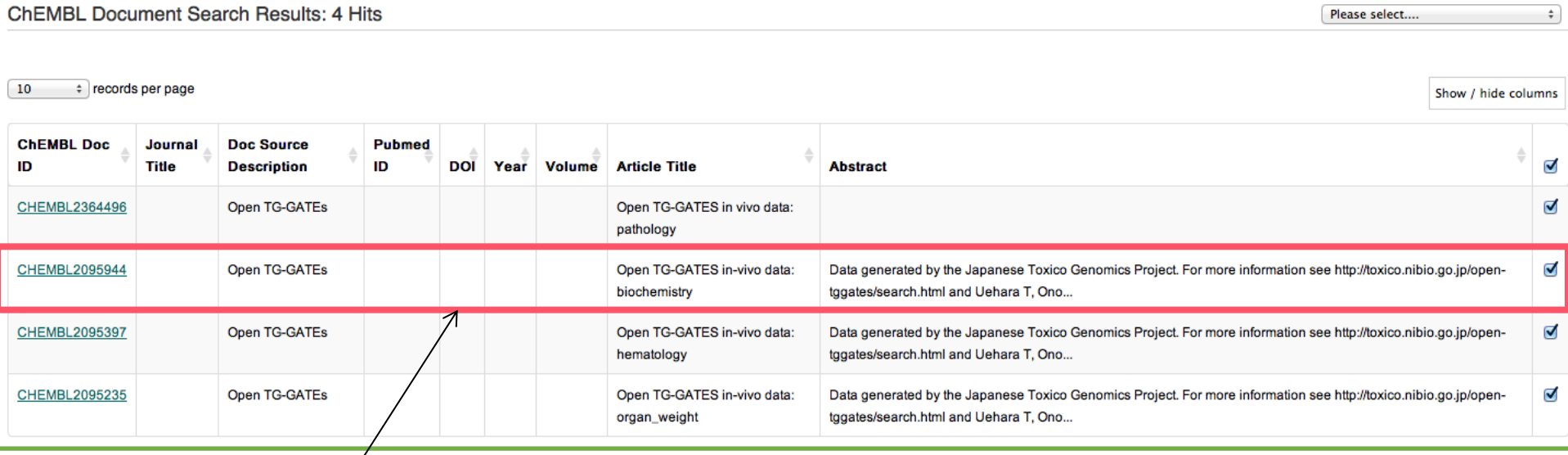
Ingredient	Molweight	Standard Type	Relation	Standard Value	Standard Units	Activity Comment	Assay Type	Description	Assay Src Description	Assay Organism	Target Type	Target Name	Target Organism	Reference
CHEMBL1275764								metabolic activation system						
 CHEMBL1275764	574.67	Activity				Toxic	A	Induction of mutagenicity in Salmonella serovar Typhimurium TA 100 assessed as number of revertant colonies at 100 ug by Ames assay	Scientific Literature	Salmonella enterica subsp. enterica serovar Typhimurium	ADMET	ADMET		Antimicrob. Agents Chemother., (2010) 54:6:2507
 CHEMBL1275764	574.67	Activity				Toxic	A	Induction of mutagenicity in Salmonella serovar Typhimurium TA 100 assessed as number of revertant colonies at 500 ug by Ames assay	Scientific Literature	Salmonella enterica subsp. enterica serovar Typhimurium	ADMET	ADMET		Antimicrob. Agents Chemother., (2010) 54:6:2507
 CHEMBL1276609	366.28	Activity				Toxic	A	Genotoxicity in Salmonella Typhimurium by Ames test	Scientific Literature	Salmonella enterica subsp. enterica serovar Typhimurium	ADMET	ADMET		Eur. J. Med. Chem., (2010) 45:11:4879
 CHEMBL458102	413.16	Activity				Toxic	A	Genotoxicity in Salmonella Typhimurium by Ames test	Scientific Literature	Salmonella enterica subsp. enterica serovar Typhimurium	ADMET	ADMET		Eur. J. Med. Chem., (2010) 45:11:4879

Results in activity comment field

Exercise 3 – Activity Source Filter



Exercise 3 –TG-GATES Datasets



ChEMBL Document Search Results: 4 Hits

Please select....

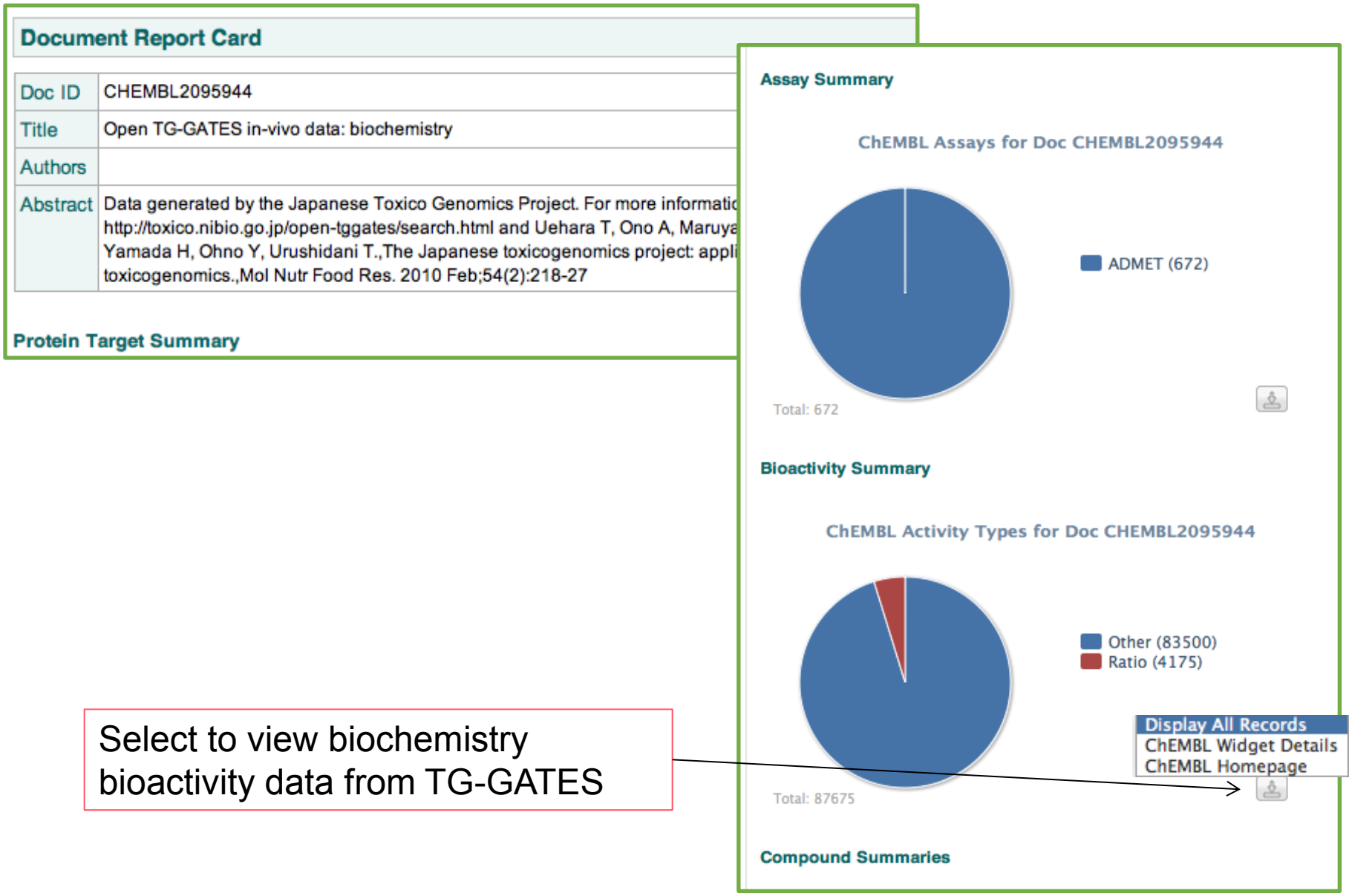
10 records per page

Show / hide columns

ChEMBL Doc ID	Journal Title	Doc Source Description	Pubmed ID	DOI	Year	Volume	Article Title	Abstract	
CHEMBL2364496		Open TG-GATEs					Open TG-GATES in vivo data: pathology		☑
CHEMBL2095944		Open TG-GATEs					Open TG-GATES in-vivo data: biochemistry	Data generated by the Japanese Toxico Genomics Project. For more information see http://toxico.nibio.go.jp/open-tggates/search.html and Uehara T, Ono...	☑
CHEMBL2095397		Open TG-GATEs					Open TG-GATES in-vivo data: hematology	Data generated by the Japanese Toxico Genomics Project. For more information see http://toxico.nibio.go.jp/open-tggates/search.html and Uehara T, Ono...	☑
CHEMBL2095235		Open TG-GATEs					Open TG-GATES in-vivo data: organ_weight	Data generated by the Japanese Toxico Genomics Project. For more information see http://toxico.nibio.go.jp/open-tggates/search.html and Uehara T, Ono...	☑

Select to view biochemistry data

Exercise 3 – TG-GATES Biochemistry Data



Select to view biochemistry bioactivity data from TG-GATES

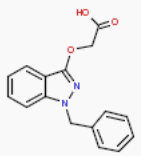
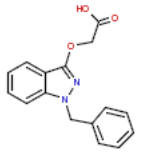
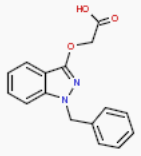
Exercise 3 – Bioactivity Results

ChEMBL Bioactivity Search Results: 83500

Please select....

100 records per page

Show / hide columns

Ingredient	Molweight	Standard Type	Relation	Standard Value	Standard Units	pChEMBL Value	Activity Comment	Assay Type	Description	Assay Src Description	Assay Organism	Target Type	Target Name	Target Organism	Reference
 CHEMBL1089221	282.29	concentration	=	6419.2	IU/L		N = 5; SD = 4954.63; p-value = 0.07	A	TG_GATES: regimen: single; time: 24 hours; dose: high | dataset: biochemistry; assay: ALP (alkaline phosphatase)	Open TG-GATEs	Rattus norvegicus	ORGANISM	Rattus norvegicus	Rattus norvegicus	CHEMBL2095944
 CHEMBL1089221	282.29	concentration	=	5387.2	IU/L		N = 5; SD = 4041.51; p-value = 0.06	A	TG_GATES: regimen: single; time: 9 hours; dose: high | dataset: biochemistry; assay: ALP (alkaline phosphatase)	Open TG-GATEs	Rattus norvegicus	ORGANISM	Rattus norvegicus	Rattus norvegicus	CHEMBL2095944
 CHEMBL1089221	282.29	concentration	=	3422.2	IU/L		N = 5; SD = 2691.45; p-value = 0.09	A	TG_GATES: regimen: single; time: 6 hours; dose: middle | dataset: biochemistry; assay: ALP (alkaline phosphatase)	Open TG-GATEs	Rattus norvegicus	ORGANISM	Rattus norvegicus	Rattus norvegicus	CHEMBL2095944

Note size of data set: 83500 rows as multiple dose, time points, end points