

CDD → BioAssay Express

Autonomous model building with a preponderance
of well annotated assay protocols

Alex M. Clark

<http://www.bioassayexpress.com>



COLLABORATIVE DRUG DISCOVERY

Background

- ◆ Assay protocols are text
- ◆ Institutions suffering from data overload:
 - ▷ much interest in using consistent markup
 - ▷ semantic web ontologies: BAO, DTO, CLO, GO...
- ◆ Since last talk: many refinements, some new structure-activity analysis features



Cell-Free Homogeneous Primary HTS to Identify Inhibitors of GSK3beta Activity

- (1) Dispense 1 uL/well of CABPE, 0.5 uL of ATP, and 1 uL of positive control GW8510 or AB in respective wells according to plate design to 1536-well assay ready plates (Aurora 29847) that contain 2.5 nL/well of 10 mM compound using BioRAPTR (Beckman) to start the reaction. Incubate at room temperature for 60 minutes.
- (2) Add 2.5 uL/well of ADP-glo (Promega, V9103) with BioRAPTR, incubate at room temperature for 40 minutes
- (3) Add 5 uL/well of ADP-glo (Promega, V9103) with Combi nL (Thermo), incubate at room temperature for 30 minutes

Curation

Assign PubChem Assay ID 651 x

www.bioassayexpress.com/assign.jsp?uniqueID=pubchemAID%3A651937

CDD BioAssay Express Assign Assay

⚡ New + Clone ↺ Random find by ID

keyword search

PubChem Assay ID 651937

Protocol Full Text

Counterscreen for discovery of small molecules that bind to the HIV-1-gp120 binding antibody, PG9: TR-FRET-based biochemical high throughput dose response assay to identify small molecules that bind to the control antibody, PGV04, which binds to a distinct site on the HIV envelope different from the PG9 binding site

The objective of this proposal is to develop and implement screening assays to utilize existing human mAbs with broad neutralizing activity (bNABs) to screen a large and diverse small molecule library to identify haptens that bind specifically to the antibody combining site of the bNAB. These small molecule mimetics of the HIV-1 envelope protein could be useful in development of an AIDS vaccine.

This project will exploit the antigen-binding property of newly discovered broad and potent neutralizing antibodies to select complementary small molecule leads that can be used as building blocks for the development of a vaccine to prevent infection with the HIV-1 virus.

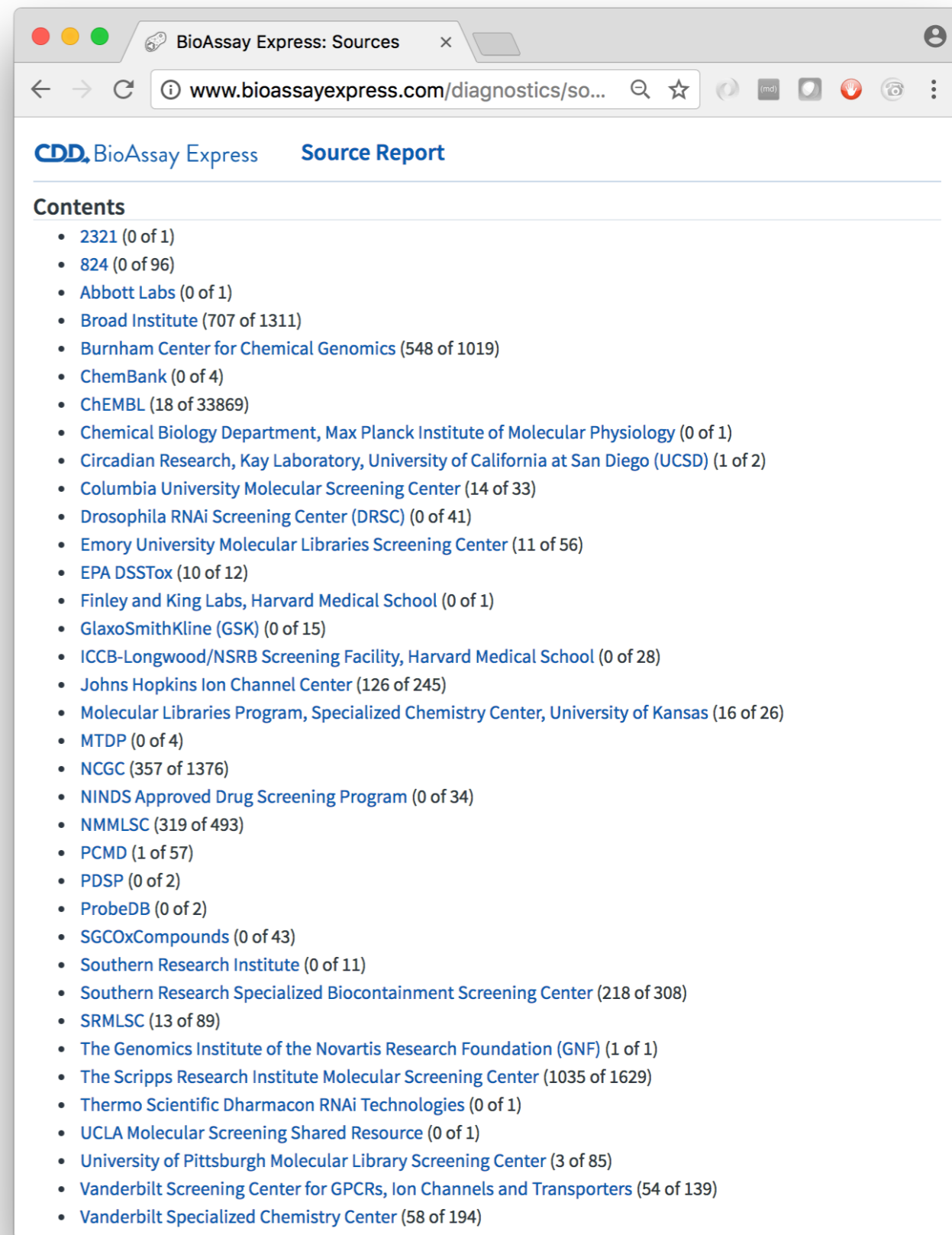
common assay template

- bioassay type > binding type x
- bioassay > protein-protein interaction assay x
- assay format > biochemical format x
- assay design method > binding assessment method x
- assay supporting method > affinity selection ✓ x
- assay cell line > not applicable x
- organism > Human immunodeficiency virus 1 x
- biological process > immunoglobulin production involved in immunoglobulin mediated immune response x
 - requires term x
- target > needs checking x
 - "HIV1: gp120" x
- applies to disease > Human immunodeficiency virus infectious disease x
- assay mode of action > inhibition x
- result > IC50 x
- screening campaign stage > alternate target assay x

- ◆ Web UI
- ◆ Machine learning
- ◆ Common Assay Template
- ◆ PubChem data

Public Data

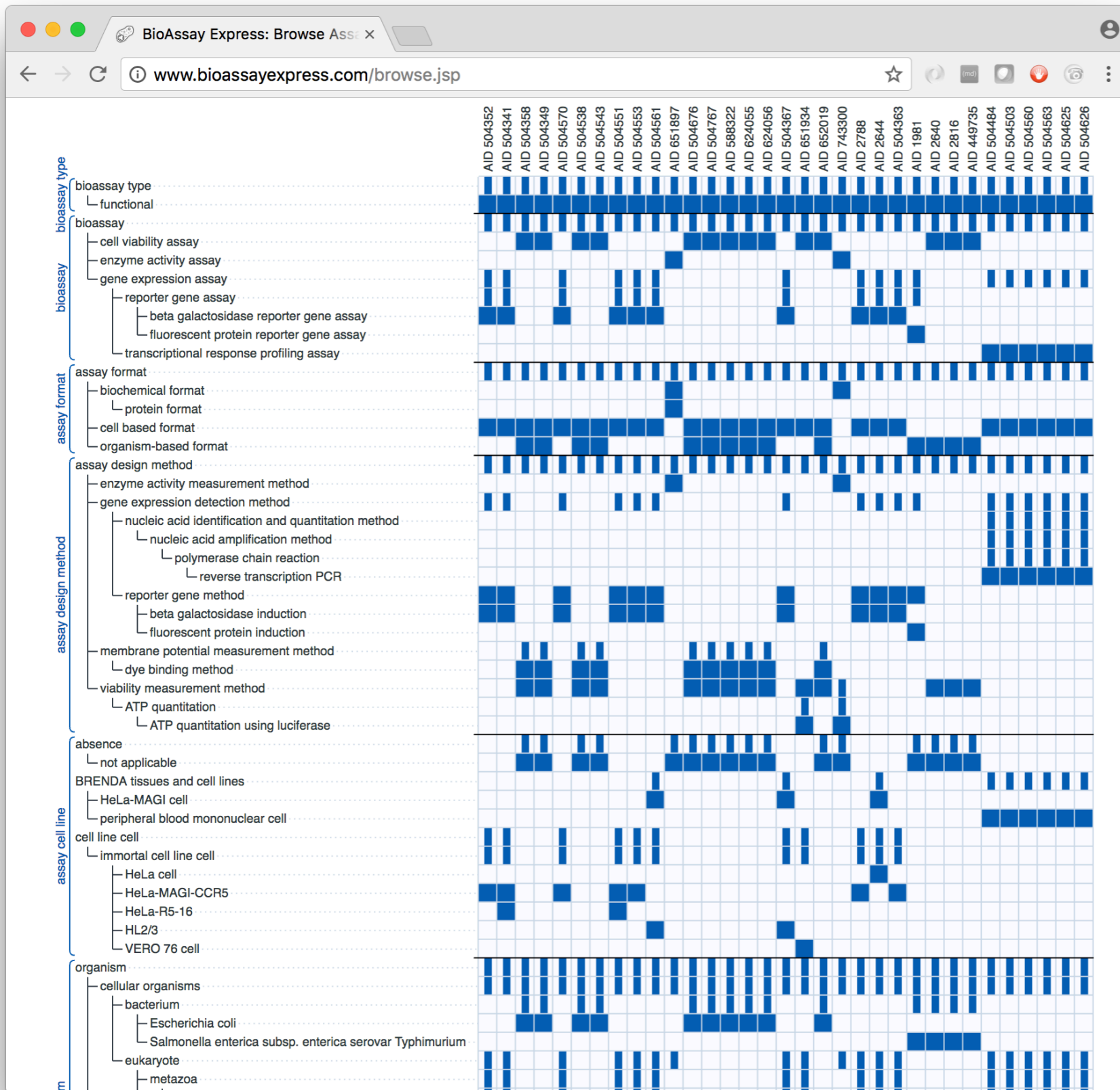
- ◆ Designed Common Assay Template for summary
- ◆ Data from PubChem:
 - ▷ 3500 assays selected
 - ▷ mostly MLPCN
- ◆ Curated as a validation set
- ◆ PubChem also has molecules...



The screenshot shows a web browser window with the URL www.bioassayexpress.com/diagnostics/so.... The page title is "CDD BioAssay Express Source Report". Under the "Contents" section, there is a list of assay sources with their respective counts:

- 2321 (0 of 1)
- 824 (0 of 96)
- Abbott Labs (0 of 1)
- Broad Institute (707 of 1311)
- Burnham Center for Chemical Genomics (548 of 1019)
- ChemBank (0 of 4)
- ChEMBL (18 of 33869)
- Chemical Biology Department, Max Planck Institute of Molecular Physiology (0 of 1)
- Circadian Research, Kay Laboratory, University of California at San Diego (UCSD) (1 of 2)
- Columbia University Molecular Screening Center (14 of 33)
- Drosophila RNAi Screening Center (DRSC) (0 of 41)
- Emory University Molecular Libraries Screening Center (11 of 56)
- EPA DSSTox (10 of 12)
- Finley and King Labs, Harvard Medical School (0 of 1)
- GlaxoSmithKline (GSK) (0 of 15)
- ICCB-Longwood/NSRB Screening Facility, Harvard Medical School (0 of 28)
- Johns Hopkins Ion Channel Center (126 of 245)
- Molecular Libraries Program, Specialized Chemistry Center, University of Kansas (16 of 26)
- MTDP (0 of 4)
- NCGC (357 of 1376)
- NINDS Approved Drug Screening Program (0 of 34)
- NMMLSC (319 of 493)
- PCMD (1 of 57)
- PDSP (0 of 2)
- ProbeDB (0 of 2)
- SGCoxCompounds (0 of 43)
- Southern Research Institute (0 of 11)
- Southern Research Specialized Biocontainment Screening Center (218 of 308)
- SRMLSC (13 of 89)
- The Genomics Institute of the Novartis Research Foundation (GNF) (1 of 1)
- The Scripps Research Institute Molecular Screening Center (1035 of 1629)
- Thermo Scientific Dharmacon RNAi Technologies (0 of 1)
- UCLA Molecular Screening Shared Resource (0 of 1)
- University of Pittsburgh Molecular Library Screening Center (3 of 85)
- Vanderbilt Screening Center for GPCRs, Ion Channels and Transporters (54 of 139)
- Vanderbilt Specialized Chemistry Center (58 of 194)

Analysis



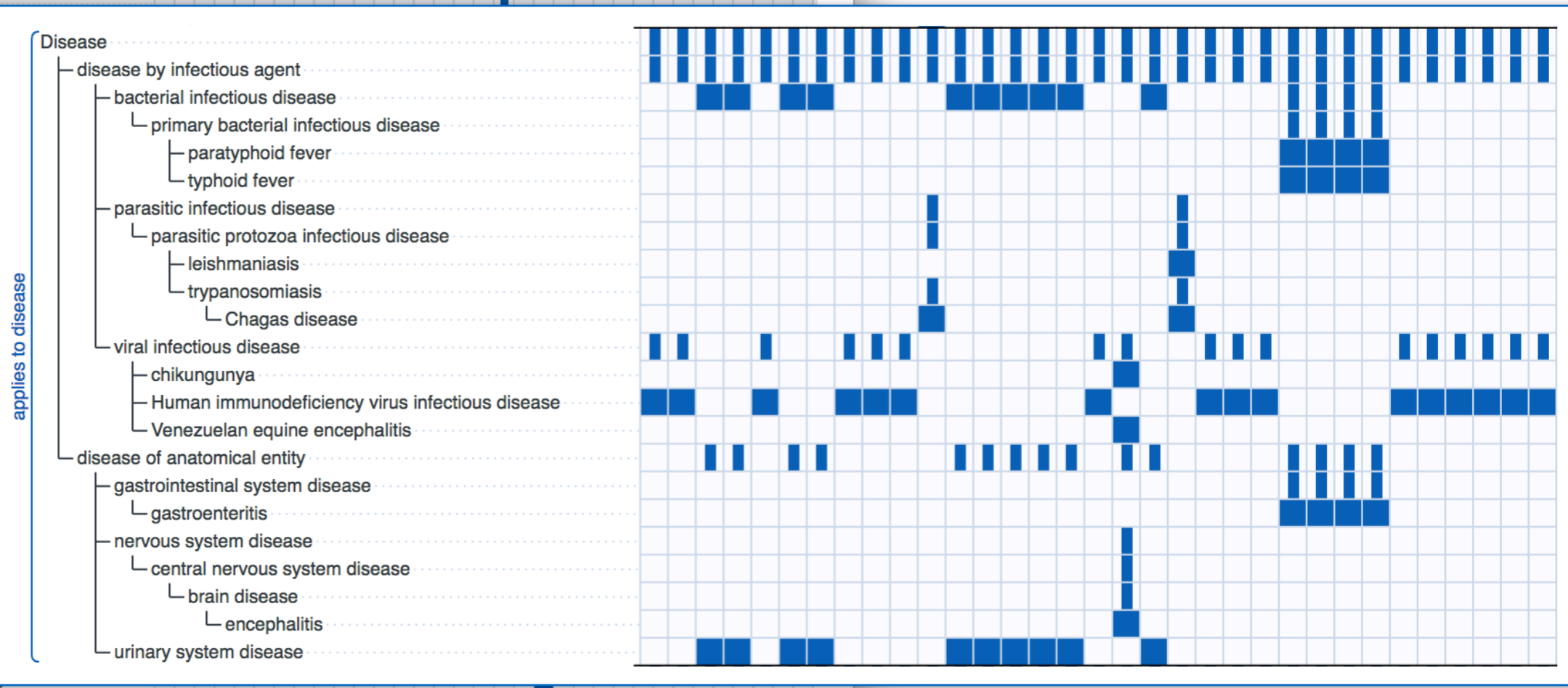
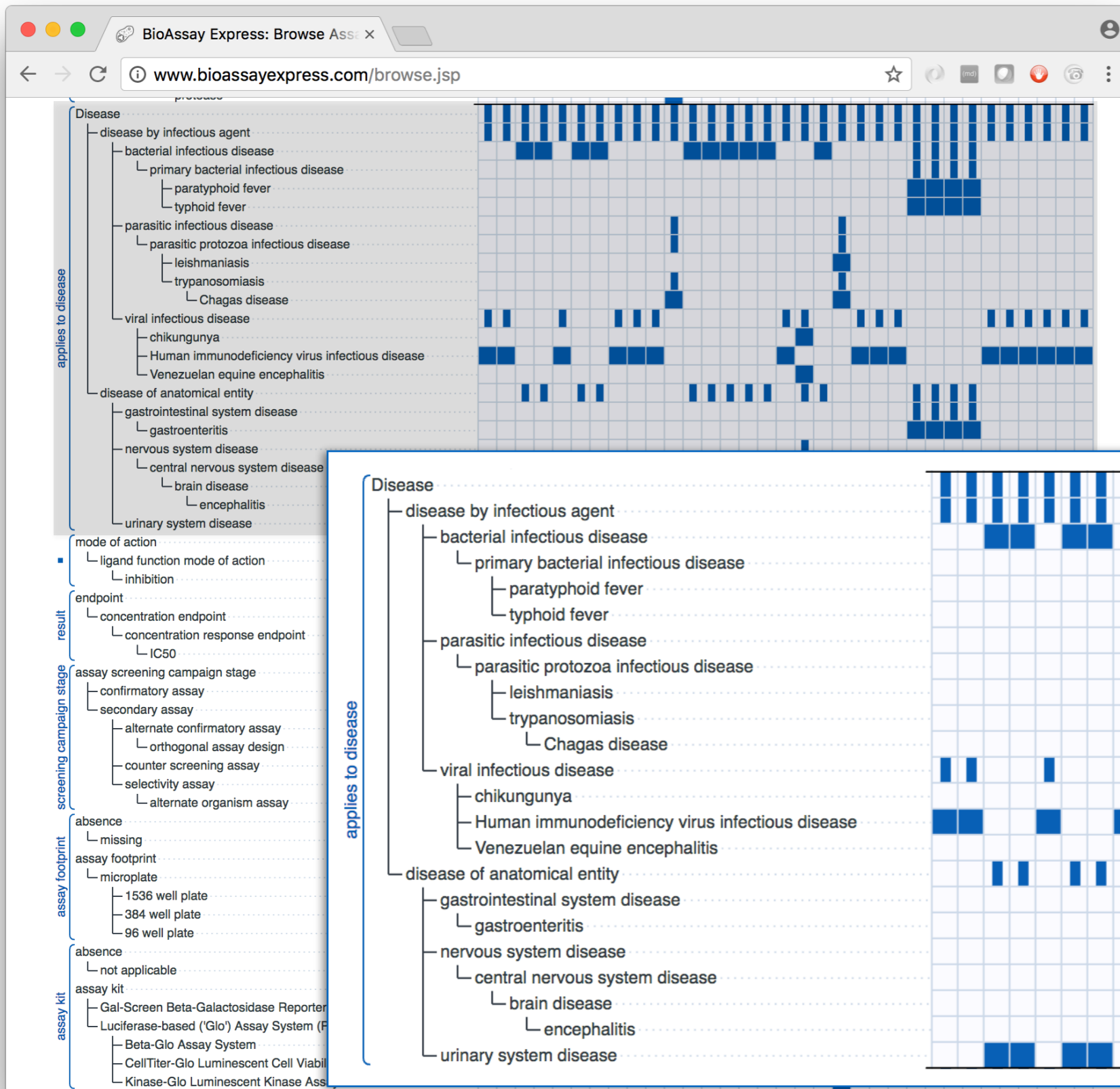
◆ Selected:

- ▷ Secondary assays
- ▷ Infectious diseases
- ▷ IC50 results

Analysis

◆ Selected:

- ▷ Secondary assays
- ▷ Infectious diseases
- ▷ IC50 results



Molecules

Assign PubChem Assay ID 651 x

www.bioassayexpress.com/assign.jsp?uniqueID=pubchemAID%3A651937

1 of 95 Show Values

3.662	2.138	7.988	3.553	2.417	2.857	3.696	4.378	3.881	
2.358	3.442	6.084	6.410	8.600	5.444	5.146	6.271	8.952	
5.341	8.447	2.499	1.546	17.23	56.77	64.32	67.07	49.15	
17.05	67.00	18.40	67.03	21.74	33.70	10.96	43.37	39.40	
67.03	28.29	57.47	67.05	54.50	28.51	20.78	17.51	11.95	

Download as SDfile

◆ Per assay: convenient way to view & download...

Groups of Assays

PRIMARY **screening campaign stage** Keywords: optional keywords

empty


- absence
 - not applicable (8)
- assay screening campaign stage
 - confirmatory assay (1106)
 - **lead optimization assay** (269)
 - preclinical development stage (5)
 - primary assay (349)
 - secondary assay (459)
 - alternate confirmatory assay (134)
 - alternate assay conditions (81)
 - alternate assay format (55)
 - alternate assay type (109)
 - orthogonal assay design (107)
 - orthogonal assay technology (33)
 - counter screening assay (693)
 - alternate target assay (250)
 - compound toxicity assay (345)
 - parental cell line assay (20)
 - physicochemical profiling assay (44)
 - compound aggregation assay (3)
 - compound fluorescence assay (27)
 - compound redox activity assay (1)
 - variant construct assay (68)
 - selectivity assay (229)
 - alternate cell line assay (44)
 - alternate organism assay (16)
 - summary assay (1)

SECONDARY **applies to disease** Keywords: optional keywords

empty

- absence
 - missing (111 of 156)
 - not applicable (246 of 252)
 - unknown (128 of 184)
- Disease
 - disease by infectious agent (35 of 45)
 - bacterial infectious disease (88 of 140)
 - primary bacterial infectious disease (6 of 32)
 - botulism (3 of 24)
 - foodborne botulism (3 of 24)
 - paratyphoid fever (13 of 23)
 - Salmonellosis (3 of 8)
 - tuberculosis (26 of 54)
 - multidrug-resistant tuberculosis (20 of 38)
 - typhoid fever (13 of 23)
 - parasitic infectious disease
 - parasitic helminthiasis infectious disease (13 of 24)
 - fascioliasis (3)
 - parasitic protozoa infectious disease
 - leishmaniasis (23 of 32)
 - malaria (2)
 - Plasmodium falciparum malaria (33 of 45)
 - trypanosomiasis (2 of 5)
 - Chagas disease (34 of 53)
 - sleeping sickness (12 of 17)
 - viral infectious disease (13 of 35)
 - chikungunya (3)
 - hepatitis C (11 of 22)

Matching assays: 366

 Compounds	Result	ID
<input checked="" type="checkbox"/> 7	Vero 76 Cytotoxicity Assay for VEEV Compounds (2)	View AID 602240
<input checked="" type="checkbox"/> 20	Inhibition of T. cruzi Tulahuen strain amastigotes in L-6 muscle cells Measured in Cell-Based and Microorganism Combination System Using Plate Reader - 2138-03_Inhibitor_Dose_CherryPick_Activityi	View AID 651881
<input checked="" type="checkbox"/> 8	L-6 rat myocyte cell toxicity Measured in Cell-Based System Using Plate Reader - 2138-06_Inhibitor_Dose_DryPowder_Activity	View AID 651895
<input checked="" type="checkbox"/> 51	A Titer Reduction Assay to evaluate Selected Compounds as new Inhibitors of Respiratory Syncytial Virus (RSV)	View AID 449732
<input checked="" type="checkbox"/> 4	A Cell-Based Counter Screen testing Compounds that Inhibit VEEV TC-83, in strain KU V3526 (2)	View AID 651884

Bayesian Models

Assay Models

Build Group By

1	AID463235	Late stage probe development assay for inhibitors of the Hepatitis C Virus non-structural protei ...
2	AID519	Dose Response Assay for Formylpeptide Receptor (FPR) Ligands and Dose Response Counter-Screen As ...
3	AID699	Secondary Calcium Response Assay for Formylpeptide Receptor (FPR) Antagonists
4	AID762	Formylpeptide Receptor (FPRL1) Ligand Structure Activity Relationship (SAR) Analysis : FPR Dose ...
5	AID763	Formylpeptide Receptor-Like-1 (FPRL1) Ligand Structure Activity Relationship (SAR) Analysis : FP ...
6	AID863	Formylpeptide Receptor (FPR) Structure Activity Relationship (SAR) Analysis: Secondary Calcium R ...
7	AID1563	Counterscreen assay for S1P4 agonists: Cell-based high throughput screening assay to identify ag ...
8	AID1650	A Cell Based Secondary Assay To Explore Cytotoxicity of West Nile Virus Anti-Viral Hit Compounds
9	AID1701	Fluorescence-based counterscreen assay for S1P4 agonists: Cell-based dose response high throughp ...
10	AID1821	Fluorescence-based counterscreen assay for S1P4 antagonists: Cell-based dose response high throu ...
11	AID1856	Epi-absorbance-based counterscreen for selective VIM-2 inhibitors: biochemical high throughput s ...
12	AID1857	FRET-based counterscreen assay for selective VIM-2 inhibitors: biochemical high throughput scree ...
13	AID1866	Epi-absorbance-based counterscreen assay for selective VIM-2 inhibitors: biochemical high throug ...
14	AID1920	Epi-absorbance-based counterscreen for selective VIM-2 inhibitors: dose response biochemical hig ...
15	AID1925	Epi-absorbance-based counterscreen for selective VIM-2 inhibitors: dose response biochemical hig ...
16	AID1926	FRET-based counterscreen for selective VIM-2 inhibitors: dose response biochemical high throughp ...
17	AID1927	FRET-based counterscreen for selective VIM-2 inhibitors: dose response biochemical high throughp ...
18	AID1944	Luminescence-based counterscreen for inhibitors of the SARS coronavirus 3C-like Protease (3CLPro ...
19	AID1981	A screen for inhibitors of the PhoP regulon in Salmonella Typhimurium using a modified countersc ...
20	AID2178	QFRET-based counterscreen for inhibitors of PFM18AAP: biochemical high throughput confirmation a ...
21	AID2184	Epi-absorbance-based counterscreen assay for common VIM-2 and IMP-1 inhibitors: biochemical high ...
22	AID2196	QFRET-based counterscreen for inhibitors of PFM18AAP: biochemical high throughput dose response ...
23	AID2244	A Cell Based Secondary Assay To Explore Cytotoxicity of West Nile Virus Anti-Viral Synthesized/A ...
24	AID2253	A cytotoxicity screen of small molecule inhibitors of the PhoP regulon in Salmonella typhimurium ...
25	AID2286	A Cell Based Assay for the Identification of Lead Compounds with Inhibitory Activity against HIV ...
26	AID2328	Counterscreen for inhibitors of EBNA-1: fluorescence polarization-based biochemical high through ...
27	AID2340	Late stage counterscreen assay for S1P4 antagonists: Fluorescence dose response cell based scree

Bayesian Models

CDD BioAssay Express

Assay Models

Build Group By

1	AID463235	Late stage probe
2	AID519	Dose Response Ass
3	AID699	Secondary Calcium
4	AID762	Formylpeptide Re
5	AID763	Formylpeptide Re
6	AID863	Formylpeptide Re
7	AID1563	Counterscreen ass
8	AID1650	A Cell Based Seco
9	AID1701	Fluorescence-bas
10	AID1821	Fluorescence-bas
11	AID1856	Epi-absorbance-b
12	AID1857	FRET-based count
13	AID1866	Epi-absorbance-b
14	AID1920	Epi-absorbance-b
15	AID1925	Epi-absorbance-b
16	AID1926	FRET-based count
17	AID1927	FRET-based count
18	AID1944	Luminescence-ba
19	AID1981	A screen for inhibi
20	AID2178	QFRET-based cou
21	AID2184	Epi-absorbance-b
22	AID2196	QFRET-based cou
23	AID2244	A Cell Based Seco
24	AID2253	A cytotoxicity scre
25	AID2286	A Cell Based Assay for the identification of Lead Compounds with inhibitory Activity against HIV ...
26	AID2328	Counterscreen for inhibitors of EBNA-1: fluorescence polarization-based biochemical high through ...
27	AID2340	Late stage counterscreen assay for S1P4 antagonists: Fluorescence dose response cell based screen

Predictions

Pick Assignments Done

- bioassay type
- bioassay
- assay format
- assay design method
- assay supporting method
- assay cell line
- organism
- biological process *bao:BAO_0002848*
- target
- applies to disease Disease targets that the assay is designed to affect. This is a more general term than the specific biological target.
- assay mode of action
- result
- screening campaign stage
- assay footprint
- assay kit
- physical detection method
- detection instrument
- perturbagen type
- protein identity
- gene identity
- GO terms
- measurement
- field
- units
- operator
- threshold

Bayesian Models

Assay Models

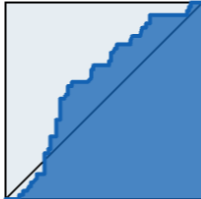


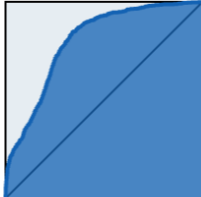
Build Group By

1	AID463235 AID623962 AID623964 AID623968 AID623973	<i>hepatitis C, hepatocellular carcinoma, viral hepatitis</i>
2	AID519 AID699 AID762 AID763 AID863	<i>bacterial infectious disease, malignant glioma</i>
3	AID1563 AID1701 AID1821 AID2349 AID2350 AID2351 AID2354 AID463107 AID463118 AID463122 AID463123 AID463129 AID463225 AID504460 AID504867 AID504869 AID504870 AID504871 AID504872 AID504873 AID504876 AID504879 AID504880 AID504881 AID504917 AID504918 AID504919 AID504921 AID504923 AID504924	<i>influenza</i>
4	AID1650 AID2244 AID2402 AID2414 AID588374 AID623930	<i>West Nile encephalitis, encephalitis</i>
5	AID1856 AID1857 AID1866 AID1920 AID1925 AID1926 AID1927 AID2184 AID2723 AID2728 AID2735 AID2755 AID2767 AID434943 AID434944 AID463099 AID463100 AID493063 AID493096 AID493097 AID493103 AID493104 AID493116 AID493117 AID493118 AID493119 AID493167 AID493172 AID504620 AID504942 AID602389 AID624011 AID624016 AID624018 AID624020 AID624033 AID624058 AID624064 AID624070 AID624072 AID624073 AID624074 AID624075 AID624080 AID624081 AID624082 AID624084 AID624085 AID624090 AID624092 AID624095 AID624096 AID624097 AID624102 AID624104 AID624105 AID624106 AID624107 AID624109 AID743441 AID743480 AID743484	<i>bacterial infectious disease</i>
6	AID1944 AID435015 AID488877	<i>severe acute respiratory syndrome</i>
7	AID1981 AID2253 AID2401 AID2640 AID2816 AID2834 AID434990 AID449735 AID449742 AID463175	<i>gastroenteritis, paratyphoid fever, typhoid fever</i>
8	AID2178 AID2196	<i>fascioliasis</i>
9	AID2286 AID2644 AID2649 AID2788 AID2812 AID435031 AID504341 AID504352 AID504363 AID504367 AID504388 AID504393 AID504394 AID504395 AID504484 AID504503 AID504530 AID504551 AID504553 AID504560 AID504561 AID504563 AID504570 AID504571 AID504573 AID504574 AID504575 AID504576 AID504625 AID504626 AID651604 AID651937 AID651939 AID652040 AID652042 AID652058 AID652062 AID687021 AID687026	<i>Human immunodeficiency virus infectious disease</i>

Bayesian Models

Assay Models

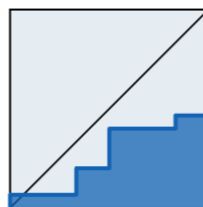
Build Group By

1	AID463235 AID623962 AID623964 AID623968 AID623973	hepatitis C, hepatocellular carcinoma, viral hepatitis		ROC 0.6094 Actives 47 of 148	Download
2	AID519 AID699 AID762 AID763 AID863	bacterial infectious disease, malignant glioma		ROC 0.8425 Actives 50 of 506	Download
3	AID1563 AID1701 AID1821 AID2349 AID2350 AID2351 AID2354 AID463107 AID463118 AID463122 AID463123 AID463129 AID463225 AID504460 AID504867 AID504869 AID504870 AID504871 AID504872 AID504873 AID504876 AID504879 AID504880 AID504881 AID504917 AID504918 AID504919 AID504921 AID504923 AID504924	influenza		ROC 0.8263 Actives 148 of 1113	Download
4	AID1650 AID2244 AID2402 AID2414 AID588374 AID623930	West Nile encephalitis, encephalitis		ROC 0.7889 Actives 409 of 933	Download
5	AID1856 AID1857 AID1866 AID1920 AID1925 AID1926 AID1927 AID2184 AID2723 AID2728 AID2735 AID2755 AID2767 AID434943 AID434944 AID463099 AID463100 AID493063 AID493096 AID493097 AID493103 AID493104 AID493116 AID493117 AID493118 AID493119 AID493167 AID493172 AID504620 AID504942 AID602389 AID624011 AID624016 AID624018 AID624020 AID624033 AID624058 AID624064 AID624070 AID624072 AID624073 AID624074 AID624075 AID624080	bacterial infectious disease		ROC 0.8005 Actives 924 of 6381	Download

Proposed Molecules

25 AID651691 AID652102

foodborne botulism



ROC 0.2667
Actives 15 of 21

[Download](#)

26 AID651738 AID651934

Venezuelan equine encephalitis, chikungunya, encephalitis



ROC 0.9750
Actives 3 of 43

[Download](#)

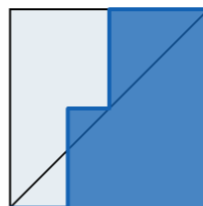
27 AID651891 AID651892

leishmaniasis

Data not suitable for model.
Actives 0 of 26

28 AID651893 AID651894

trypanosomiasis



ROC 0.5625
Actives 2 of 26

[Download](#)

29 AID743070 AID743071 AID743072 AID743073 AID743349 AID743363

tuberculosis

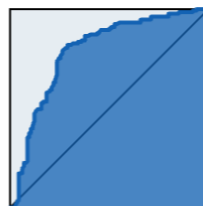


ROC 0.8733
Actives 149 of 334

[Download](#)

30 AID977615 AID977616 AID977617 AID1053115 AID1117271 AID1117272
AID1117273 AID1117275 AID1117276 AID1117284

smallpox



ROC 0.7882
Actives 133 of 247

[Download](#)

Molecules

Drag/paste molecules,



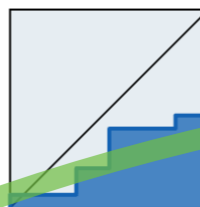
or



Proposed Molecules

25 AID651691 AID652102

foodborne botulism

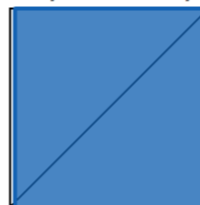


ROC 0.2667
Actives 15 of 21

Download

26 AID651738 AID651934

Venezuelan equine encephalitis, chikungunya, encephalitis

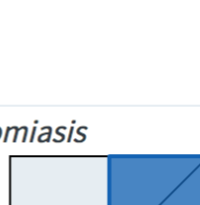


ROC 0.9750
Actives 3 of 43

Download

27 AID651891 AID651892

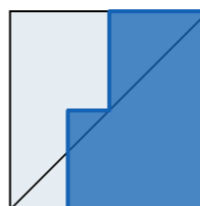
leishmaniasis



Data not suitable for model
Actives 0 of 26

28 AID651893 AID651894

trypanosomiasis



ROC 0.5625
Actives 2 of 26

Download

29 AID743070 AID743071 AID743072 AID743073 AID743349 AID743363

tuberculosis

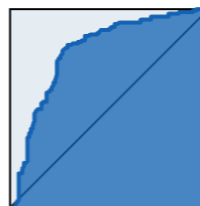


ROC 0.8733
Actives 149 of 334

Download

30 AID977615 AID977616 AID977617 AID1053115 AID1117271 AID1117272
AID1117273 AID1117275 AID1117276 AID1117284

smallpox



ROC 0.7882
Actives 133 of 247

Download

Molecules

Drag/paste molecules,

Draw

or

Example

Molecular Datasheet - forprediction.ds		
Header		
Aspects		
New View		
Export		
Graphics		
	Molecule	Name
1		aspirin
2		2-methylpyridine
3		
4		
5		
6		
7		
8		

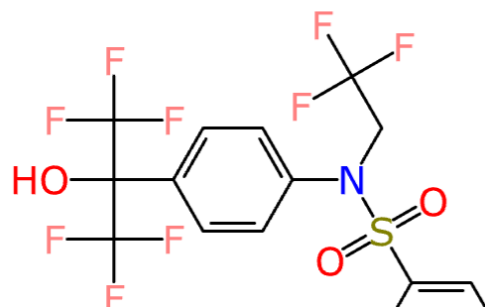
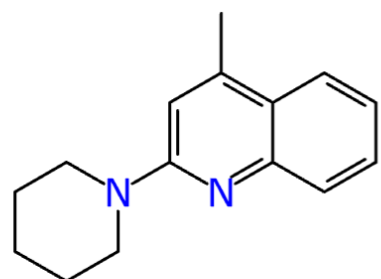
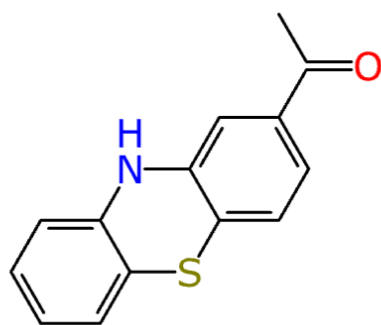
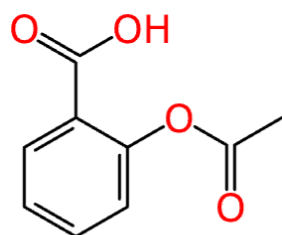
Proposed Molecules

Molecules

Drag/paste molecules,  Draw or  Example

Structure

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 28 29 30

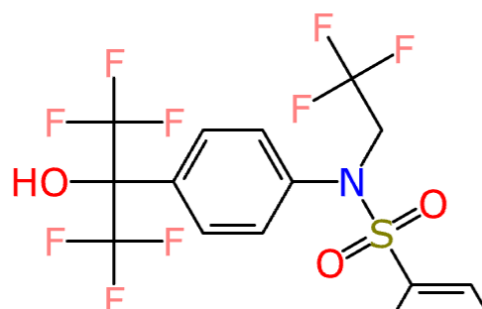
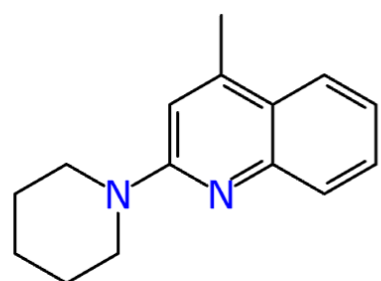
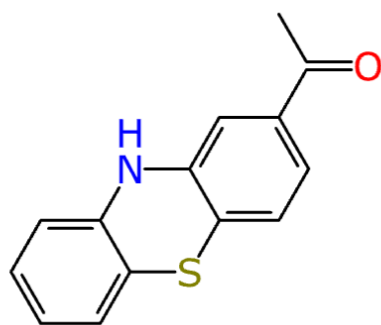
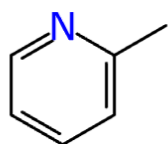
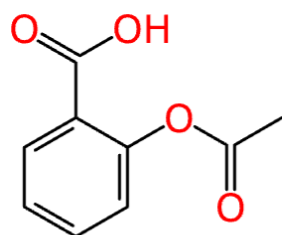


Proposed Molecules

Molecules

Drag/paste molecules,  Draw or  Example

Structure



1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 28 29 30

Model 21

Venezuelan equine encephalitis,
encephalitis, respiratory syncytial virus
infectious disease

Bayesian prediction: 91.2%



Iterative Structures

ACTIVES 2 OF 20

[Download](#)

29 AID743070 AID743071 AID743072 AID743073 AID743349 AID743363

tuberculosis



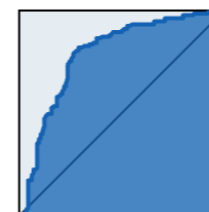
ROC 0.8733

Actives 149 of 334

[Download](#)

30 AID977615 AID977616 AID977617 AID1053115 AID1117271 AID1117272
AID1117273 AID1117275 AID1117276 AID1117284

smallpox



ROC 0.7882

Actives 133 of 247

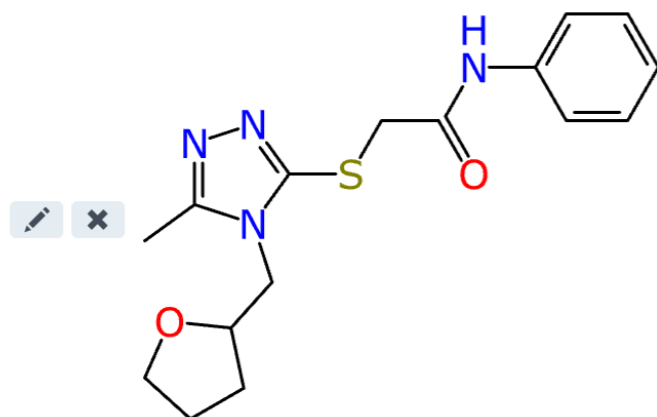
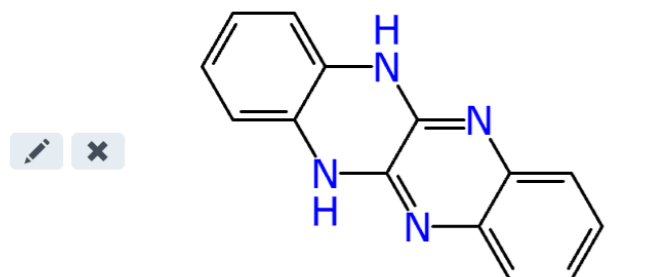
[Download](#)

Molecules

Drag/paste molecules, [Draw](#) or [Example](#)

Structure

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 28 29 30



Iterative Structures

ACTIVES 2 OF 20

Download

29 AID743070 AID743071 AID743072

30 AID977615 AID977616 AID977617
AID1117273 AID1117275 AID1117276

Edit Compound Clear Copy Close Save

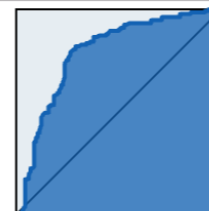
Molecules
Drag/paste molecules, Draw or

Structure

Cc1nc2c(ncn2C1CC3CCOC3)SCC(=O)Nc4ccc5c(c4)nc6c5ncn6

Iterative Structures

AID1117273 AID1117275 AID1117276 AID1117284



ROC 0.7882

Actives 133 of 247

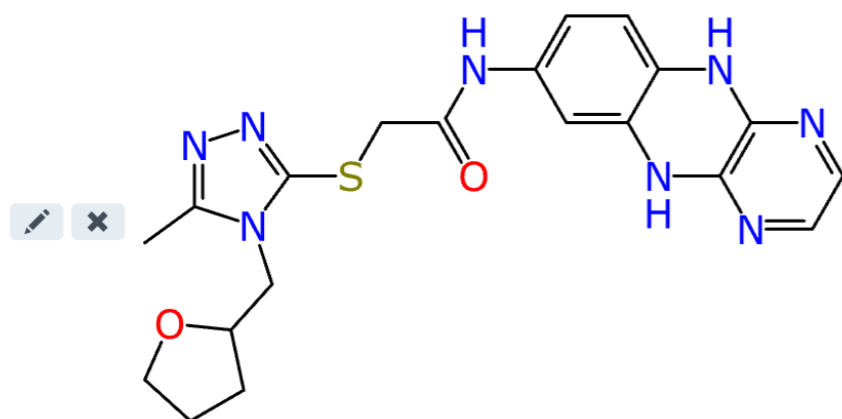
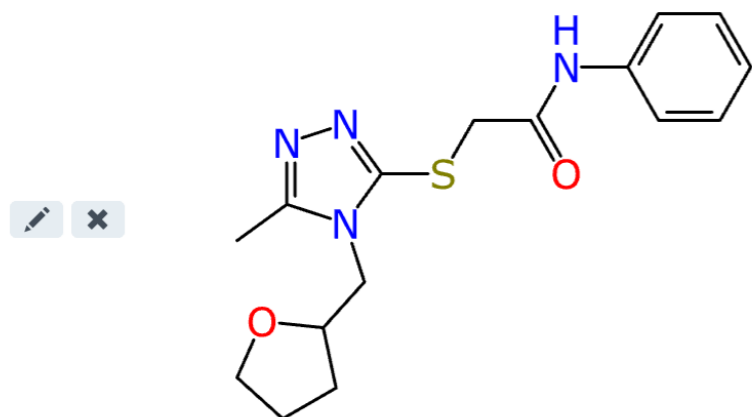
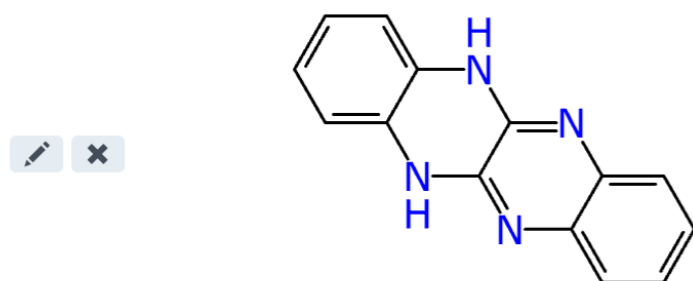
[Download](#)

Molecules

Drag/paste molecules, [Draw](#) or [Example](#)

Structure

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 28 29 30



Model 10

Burkitt lymphoma, Hodgkin's lymphoma, infectious mononucleosis, lupus erythematosus, multiple sclerosis

Bayesian prediction: >100%

Grids

- ◆ View known & calculated measurements, plotting assays vs. compounds...

CDD BioAssay Express Assay Grid

Assays

Query: 🔍 Search

Assays: 348

Compounds

Frequent Hitters Selectivity Query Structures

Maximum compounds: Few Some Many

🔍 Search

Grid

Select compounds to go with the assays, then build the grid.

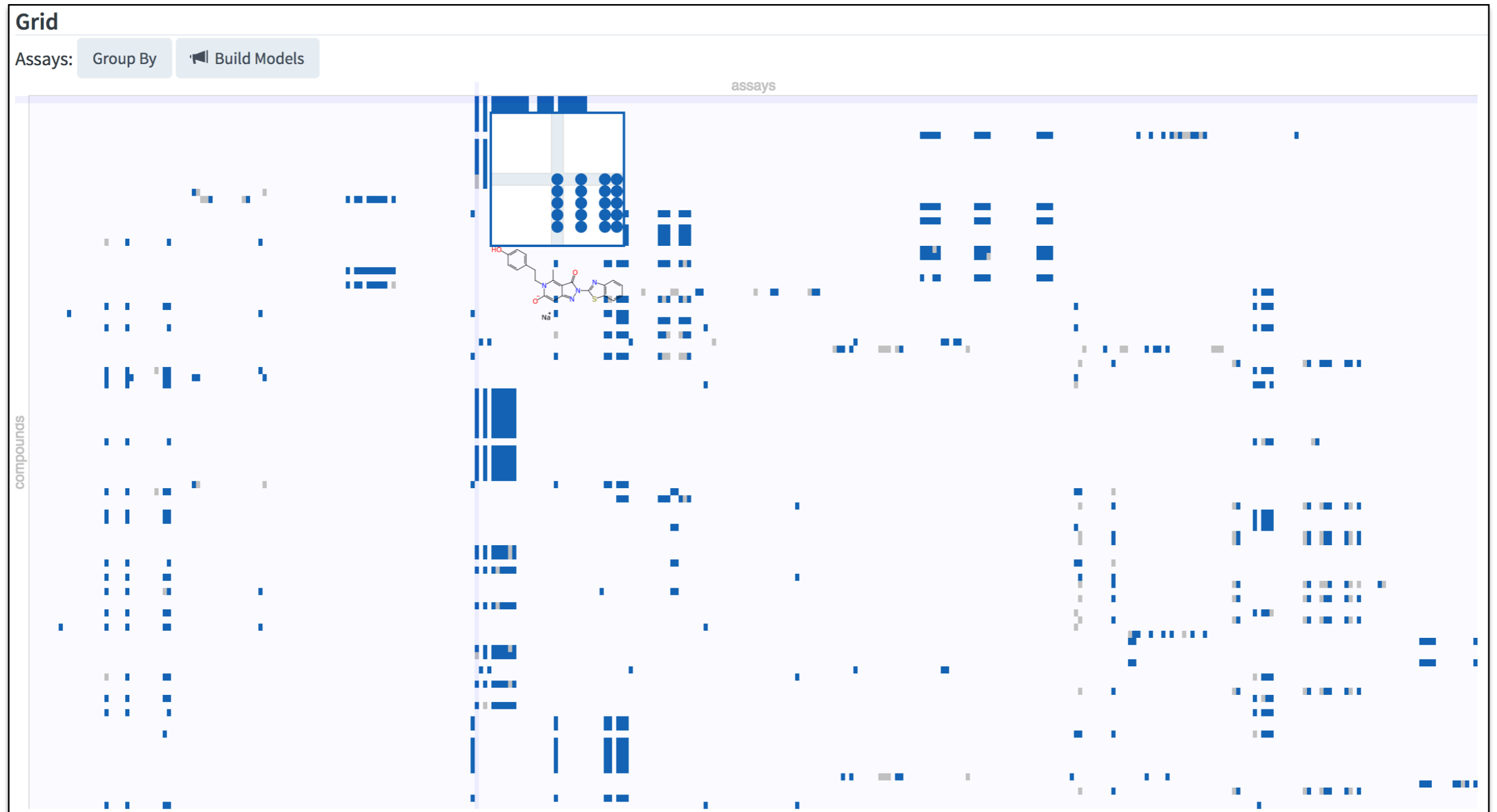
Grids

- ◆ View known & calculated measurements, plotting assays vs. compounds...

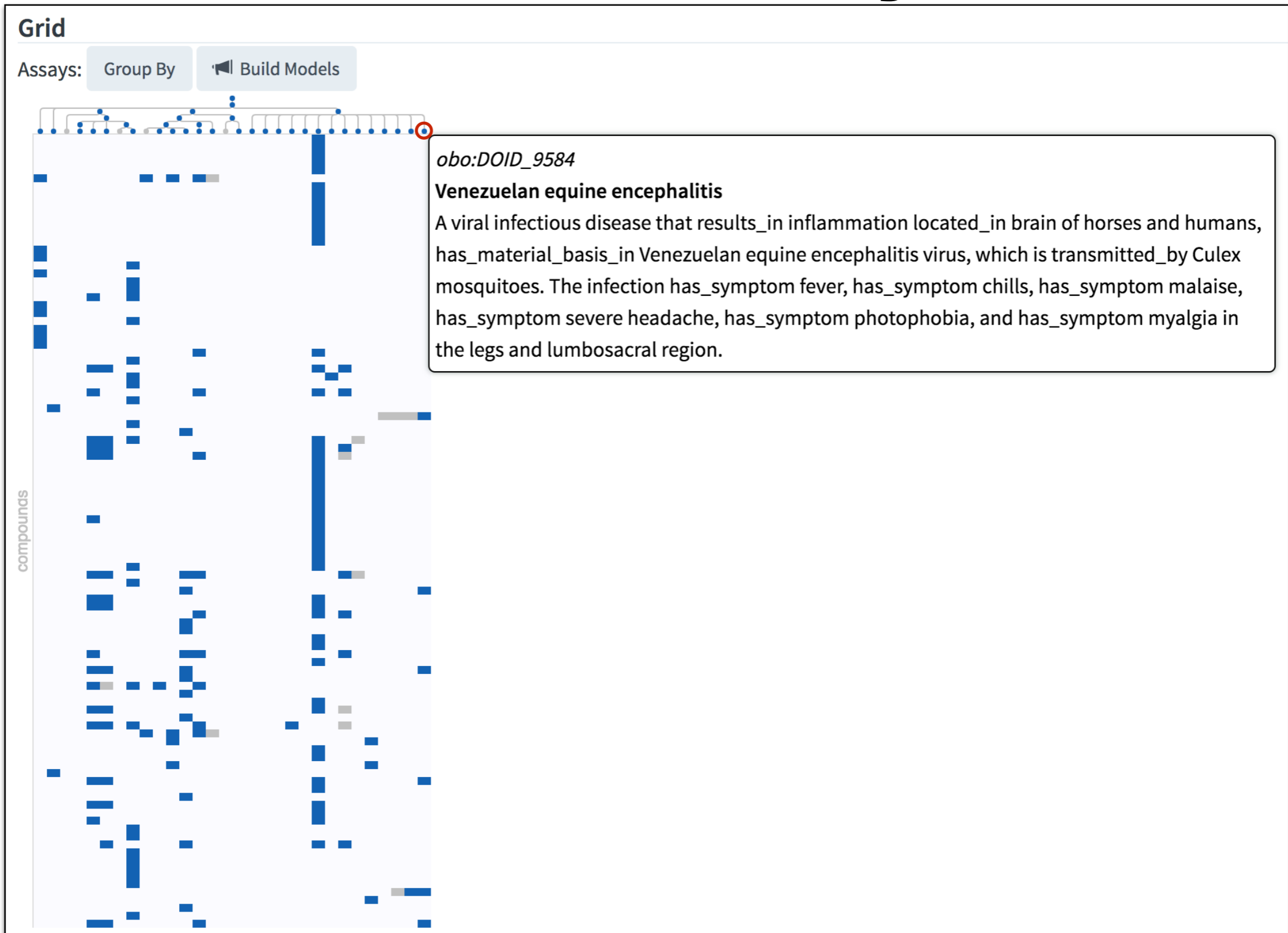
The screenshot shows the 'Assay Grid' interface in CDD BioAssay Express. At the top left is the logo 'CDD BioAssay Express' and the title 'Assay Grid'. Below this is a section for 'Assays' with a search bar labeled 'Query:' and a 'Search' button. It indicates 'Assays: 348'. The 'Compounds' section has three tabs: 'Frequent Hitters' (selected), 'Selectivity', and 'Query Structures'. Below the tabs are filters for 'Maximum compounds:' with options 'Few', 'Some' (selected), and 'Many'. A 'Stop' button is visible. A progress bar shows a blue segment on the left and a grey segment on the right, with a histogram below it. The text '5646 compounds' is displayed below the progress bar. At the bottom, the 'Grid' section contains the instruction: 'Select compounds to go with the assays, then build the grid.'

Real Data Only

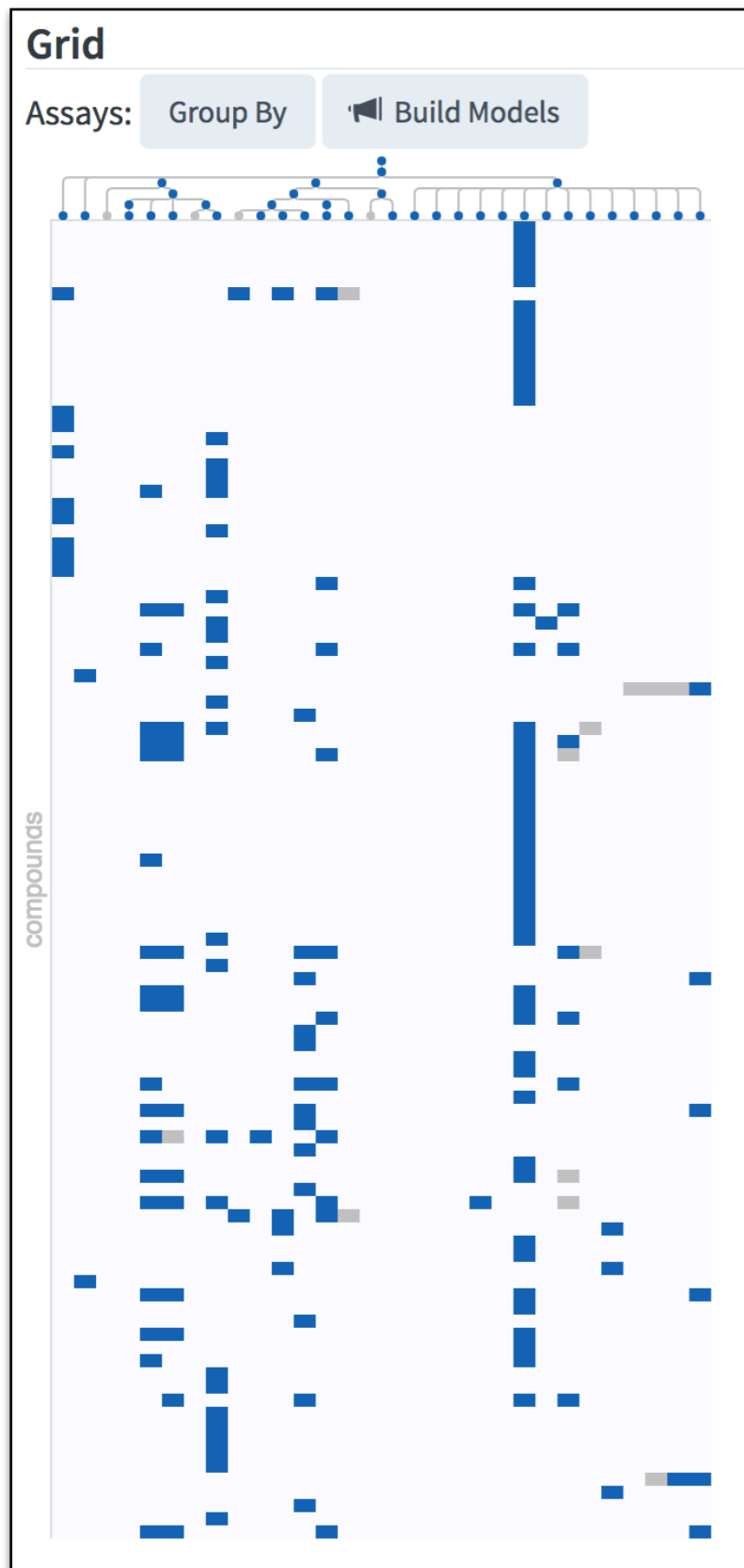
- ◆ Assays are singletons by default, can be grouped...



Real Data Only

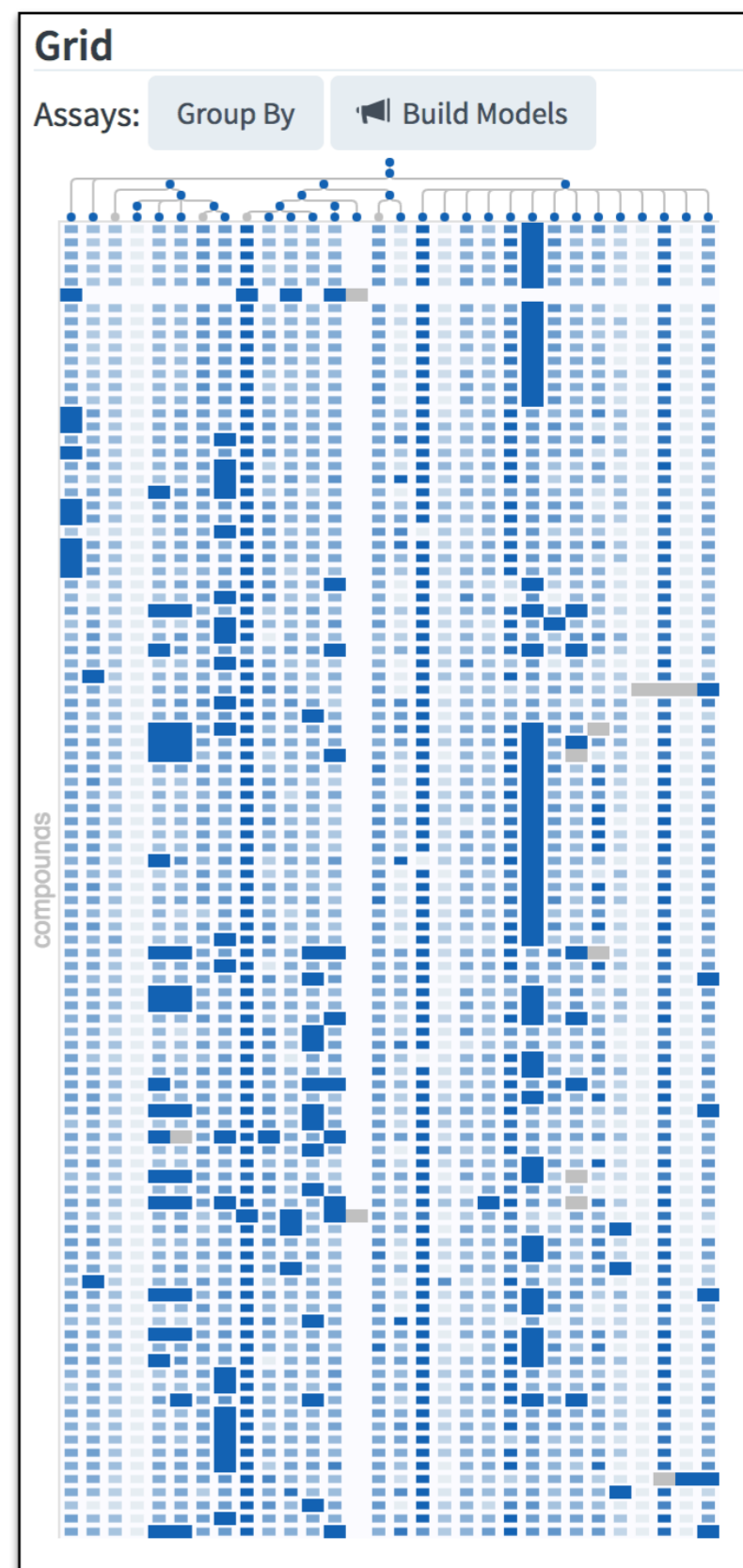


Filling in Blanks



Filling in Blanks

- ◆ Bayesian model created for each column
- ◆ SAR for all assays used
- ◆ Missing cells replaced with predictions
- ◆ Calculations realtime:
 - ▷ model on server
 - ▷ predictions on client



Trend Detection

- ◆ Model compound found using the **Selectivity** option:

Compounds

Frequent Hitters Selectivity **Query Structures**

Only select probe compounds

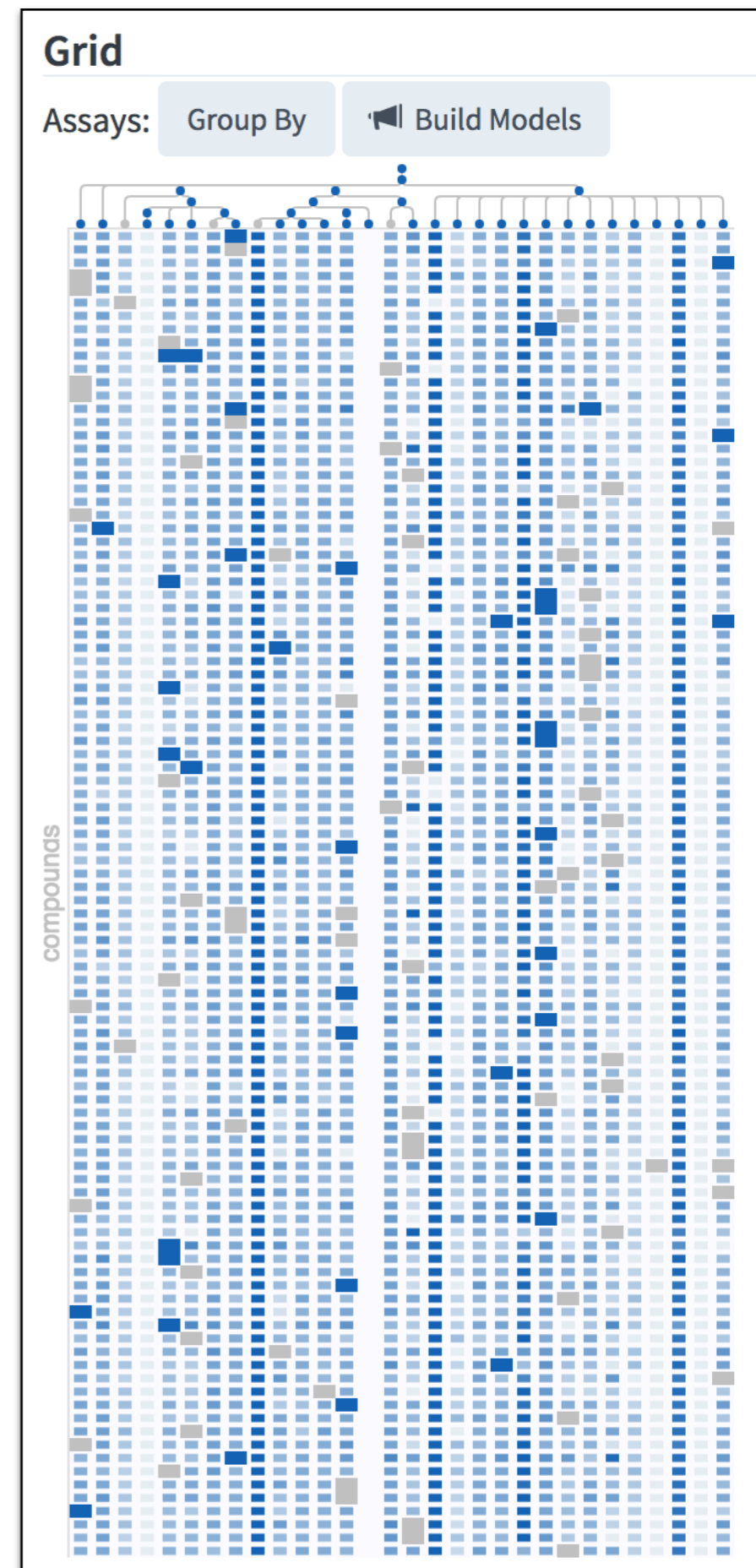
Rank by similarity to:

CCOC1=CN(C(=O)O1)C2=CC=C(C)C=C2

[Edit](#)

Maximum compounds: Few **Some** Many

Compounds found: 100



Trend Detection

- ◆ Model compound found using the **Selectivity** option:

Compounds

Frequent Hitters Selectivity **Query Structures**

Only select probe compounds

Rank by similarity to:

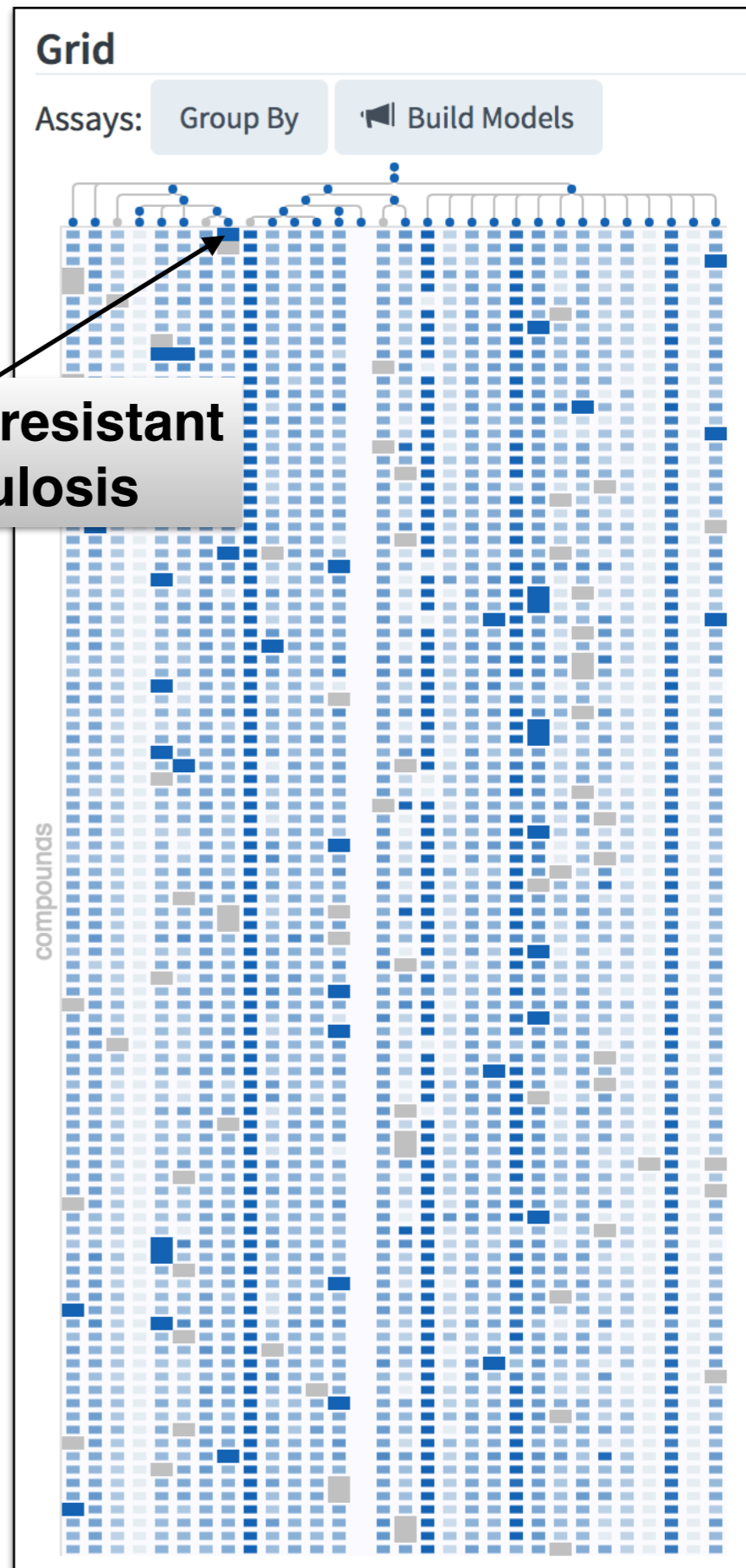
CCOC1=CN(C(=O)N)C(Cl)=C1c2ccc(C)cc2

[Edit](#)

Maximum compounds: Few **Some** Many

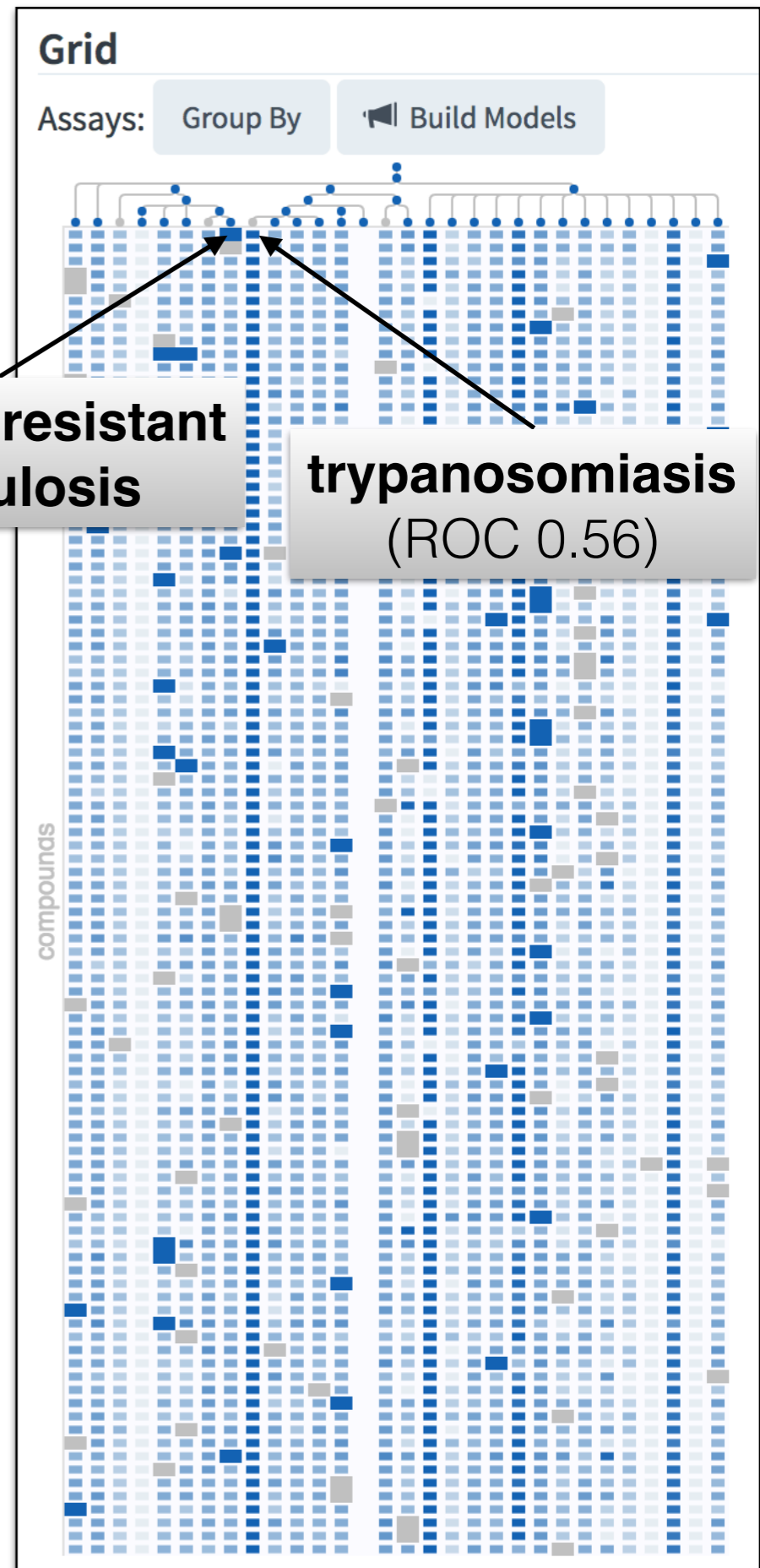
Compounds found: 100

multidrug resistant tuberculosis



Trend Detection

- ◆ Model compound found using the **Selectivity** option:



Compounds

Frequent Hitters Selectivity **Query Structures**

Only select probe compounds

Rank by similarity to:

[Edit](#)

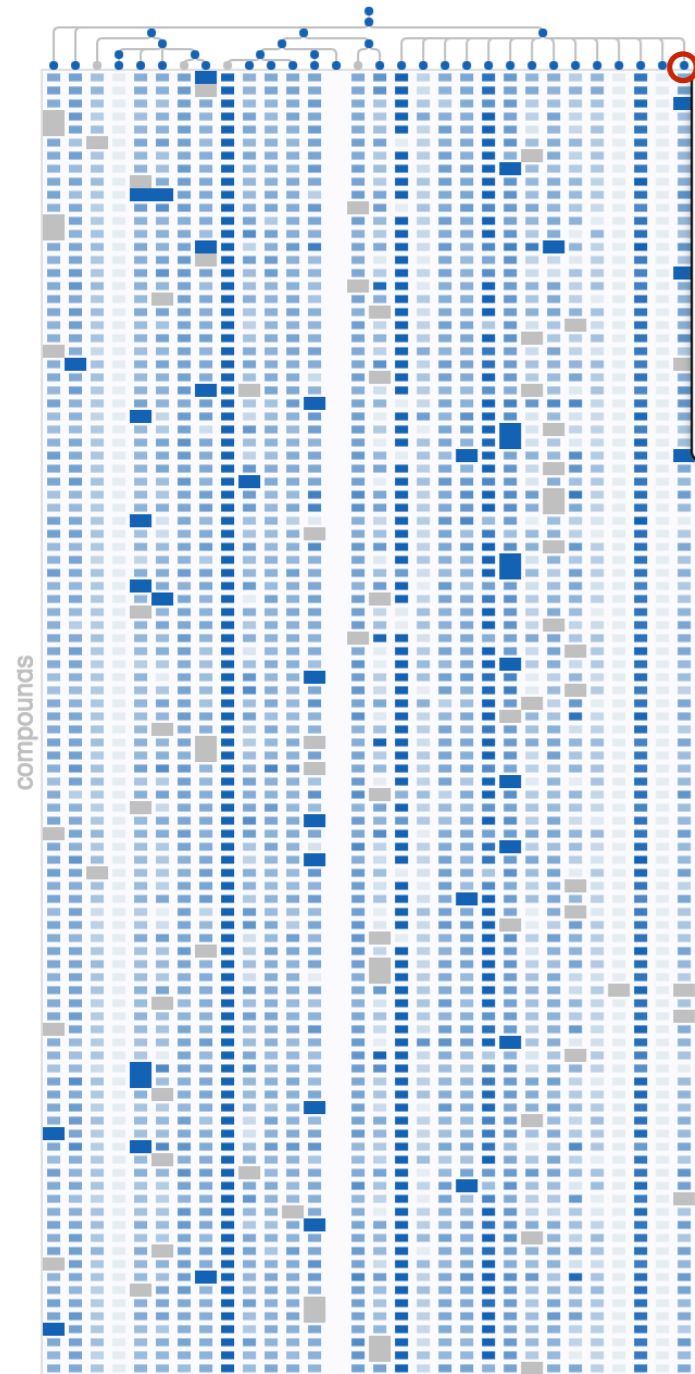
Maximum compounds: Few **Some** Many

Compounds found: 100

Possible Interest

Grid

Assays: Group By Build Models



obo:DOID_9584

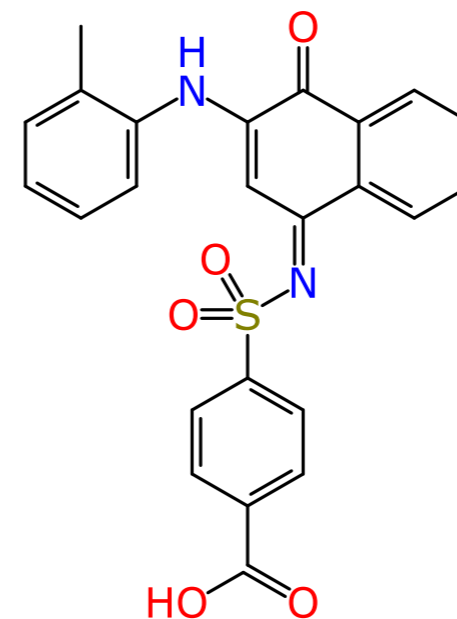
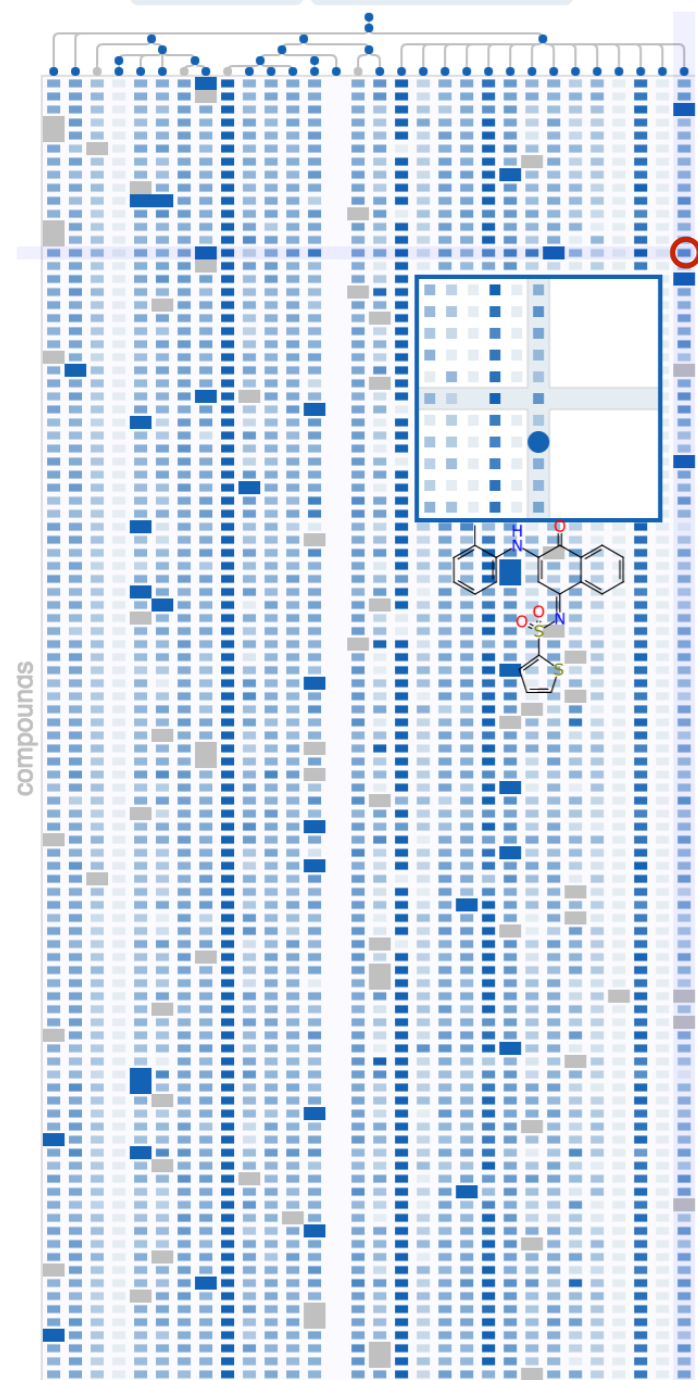
Venezuelan equine encephalitis

A viral infectious disease that results_in inflammation located_in brain of horses and humans, has_material_basis_in Venezuelan equine encephalitis virus, which is transmitted_by Culex mosquitoes. The infection has_symptom fever, has_symptom chills, has_symptom malaise, has_symptom severe headache, has_symptom photophobia, and has_symptom myalgia in the legs and lumbosacral region.

Possible Interest

Grid

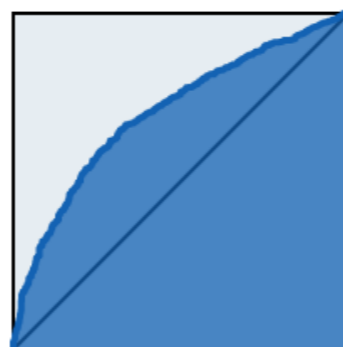
Assays: Group By Build Models



Bayesian Model

Prediction: 63.0%

Model created from 992 actives of 1660 molecules. ROC = 0.7048



Group By encephalitis
Venezuelan equine encephalitis

Column. 18 assays have: 3 × active, 4 × inactive. Row. Compound affects assays: 2 × active.

Future Work

- ◆ Current focus is UI/UX and proof-of-concept...
 - ▷ intensive data mining is next: **annotations + SAR**
 - ▷ data is ready to reveal some amazing discoveries
- ◆ Crowd curation may be necessary, and more detailed templates
- ◆ More modelling techniques, far beyond Bayesians: parallel project is underway...

Deployment

◆ Public

- ▷ <http://www.bioassayexpress.com>
- ▷ <http://beta.bioassayexpress.com> (curation)

◆ Vault

- ▷ sync with CDD Vault through API

◆ Private

- ▷ custom integration, installed behind firewall

Acknowledgments

- ◆ Collaborative Drug Discovery

 - ⇒ Barry Bunin

 - ⇒ Janice Kranz

 - ⇒ Hande Küçük

 - ⇒ Peter Gedeck

- ◆ And we're hiring

 - ⇒ resumes to:

alex@collaborativedrug.com

- ◆ More information

<http://www.bioassayexpress.com>

<http://github.com/cdd/bioassay-template>

<http://collaborativedrug.com>