

CDD BioAssay Express

Expanding the target dimension:
How to visualize a lot of SAR

Alex M. Clark

<http://www.bioassayexpress.com>



COLLABORATIVE DRUG DISCOVERY

Protocols & Datasets

◆ Structure-activity data usually exists in isolation:

- ▷ one **target**, one **protocol**
- ▷ many **compounds** & **measurements**

◆ Datasets often related:

- ▷ same target?
- ▷ same protocol?

◆ *How would we know?*

Molecule	Name	Activity

Molecule	Name	Activity
		= 59 nM
		= 90 nM
		= 170 nM
		= 25 nM
		= 11 nM

Prerequisite

- ◆ Protocols normally described using scientific English
- ◆ Alternatively, many ontologies are available:
 - ▷ **BAO** (BioAssay Ontology)
 - ▷ **DTO** (Drug Target Ontology)
 - ▷ **CLO** (Cell Line Ontology)
 - ▷ **GO** (Gene Ontology)
 - ▷ ... and many others
- ◆ Each term has a URI: global meaning, hierarchical

Common Assay Template

- ◆ Pulls together ontology categories
- ◆ ~22 assignments
- ◆ Captures key summary information

common assay template

bioassay type	>	<input type="text"/>
assay format	>	<input type="text"/>
assay design method	>	<input type="text"/>
assay supporting method	>	<input type="text"/>
assay cell line	>	<input type="text"/>
organism	>	<input type="text"/>
biological process	>	<input type="text"/>
target	>	<input type="text"/>
applies to disease	>	<input type="text"/>
assay mode of action	>	<input type="text"/>
result	>	<input type="text"/>
measurement unit	>	<input type="text"/>
screening campaign stage	>	<input type="text"/>
assay footprint	>	<input type="text"/>
assay kit	>	<input type="text"/>
physical detection method	>	<input type="text"/>
detection instrument	>	<input type="text"/>
perturbagen type	>	<input type="text"/>
protein identity	>	<input type="text"/>
gene identity	>	<input type="text"/>
GO terms	>	<input type="text"/>

Common Assay Template

- ◆ Pulls together ontology categories
- ◆ ~22 assignments
- ◆ Captures key summary information

PubChem Assay ID 720536

Protocol Full Text

TGF- β is a main component in the TGF- β signaling pathway which plays diverse roles in cellular and development pathways. TGF- β is mediated by the transcription factors Smads. Although the Smad-dependent pathway is the primary canonical TGF- β signaling node, TGF- β 1 can also activate alternative signaling pathways, including those involving MAPK (ERK, JNK and p38). This interaction may mediate or enhance Smad-dependent responses, or can exert Smad-independent effects. The complexity of this signaling cascade allows the TGF- β superfamily to perform unique, overlapping or redundant functions. We believe that targeting the TGF- β pathway at the Smad-transcription factor level may eliminate the consequences of disrupting the entire pathway and offer specificity without affecting other signaling pathways. Smad3 is the primary transducer of TGF- β 's signals and Smad3 regulates many functions attributed to TGF- β signaling. We hypothesize that Smad3 inhibitors will selectively eliminate Smad3-specific TGF- β signals without undesired off-target effects. We aim to identify Smad3-small molecule antagonists using a quantitative high throughput screening (qHTS) approach.

A high-throughput assay was developed to screen the NIH Molecular Libraries Small Molecule Repository (MLSMR). This is a cell based assay, where TGF- β is tagged with GFP. This orthogonal assay was run using CCL64 cells (as opposed to HepG2); activity was desired in this assay as this indicated the hits

Update Suggestions Stop

common assay template

bioassay type functional x
fluorescent protein reporter gene assay x

assay format cell based format x

assay design method fluorescent protein induction x

assay supporting method perturbagen delivery method (62.65%) ✓ x

assay cell line Mv 1 Lu cell x

organism Homo sapiens x

biological process SMAD protein signal transduction x
regulation of signal transduction x
sequestering of TGFbeta from receptor via TGFbeta binding x

target see Gene ID x

applies to disease missing x

assay mode of action inhibition x

result AC50 x

measurement unit micromolar x

screening campaign stage orthogonal assay design x
alternate cell line assay x

assay footprint 1536 well plate x

assay kit not applicable x

physical detection method fluorescence intensity x

detection instrument Acumen x

perturbagen type MLSMR library x

protein identity Smad3 [Homo sapiens] x

gene identity SMAD3: SMAD family member 3 x

GO terms chromatin (99.90%) ✓ x

Protocol Similarity

504854

A Cell Based Secondary Assay To Explore Vero Cell Cytotoxicity of Purified and Synthesized Compounds that Inhibit Mycobacterium Tuberculosis (4)

This functional assay was developed for detection of compounds inhibiting Vero E6 cells viability as a secondary screen to the beta-lactam sensitizing M. tuberculosis bacteriocidal assay.

In this assay, we treated Vero E6 cells with compounds selected as hits in the M. tuberculosis assay for 72 hours over a 10 point 2-fold dilution series, ranging from 0.195 uM to 100 uM. Following 72 hours of treatment, relative viable cell number was determined using Cell Titer Glo from Promega. Each plate contained 64 replicates of vehicle treated cells which served as negative controls.

Outcome: Compounds that showed <70% cell viability for at least one concentration were defined as "Active". If the % viability at all doses was >70%, the compound was defined as "Inactive".

...

69%
similar (?)

449764

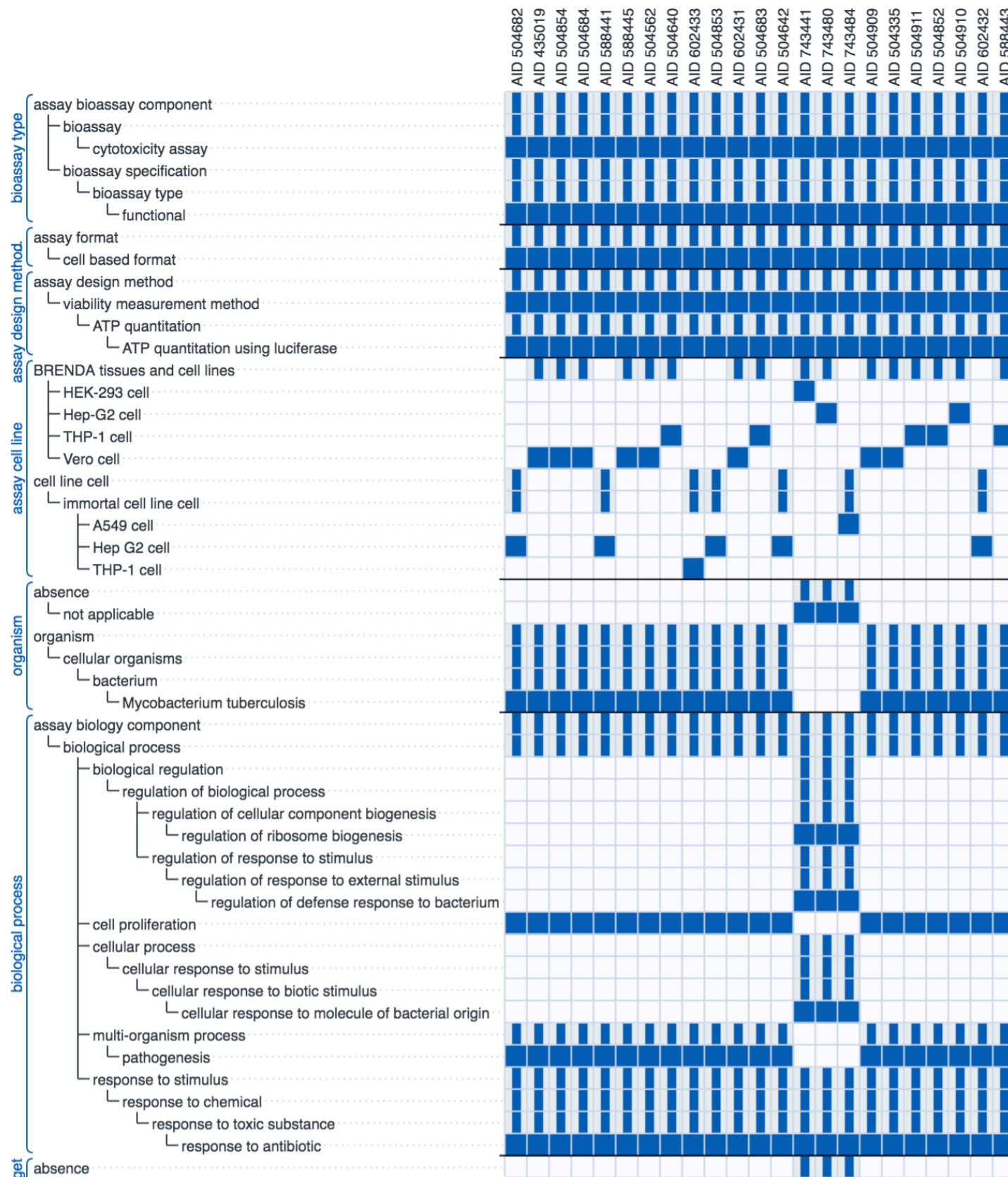
A High Throughput Confirmatory Assay used to Identify Novel Compounds that Inhibit Mycobacterium Tuberculosis in the absence of Glycerol

Outcome: Compounds that showed >30% inhibition for at least one concentration in the Mtb dose response were defined as "Active". If the inhibition at all doses was <30% in the Mtb assay, the compound was defined as "Inactive". In the primary screen a compound was deemed "Inactive" if it had a Percent Inhibition <70.31%. Compounds with a Percent Inhibition >70.31% but were not selected for follow up dose response were labeled "Inconclusive."

The following tiered system has been implemented at Southern Research Institute for use with the PubChem Score. Compounds in the primary screen are scored on a scale of 0-40 based on inhibitory activity where a score of 40 corresponds to 100% inhibition. In the confirmatory dose response screen, active compounds were scored on a scale of 41-80 based on the IC50 result in the Mtb assay while compounds that did not confirm as actives were given the score 0.

...

Protocol Similarity



semantic
annotations

≡

fingerprints

Matrix: Assays vs. Compounds

- ◆ Curated ~3500 assays from **PubChem**:
 - ▷ mostly from *Molecular Libraries* project
 - ▷ natural language/machine learning helped
 - ▷ carefully analyzed by expert biologists
- ◆ Enough data for proof of concept *assay informatics*
- ◆ Adding the **assay** dimension to **structure-activity relationship** analysis

Now With Compounds

◆ Each assay has compounds & data from **PubChem**:

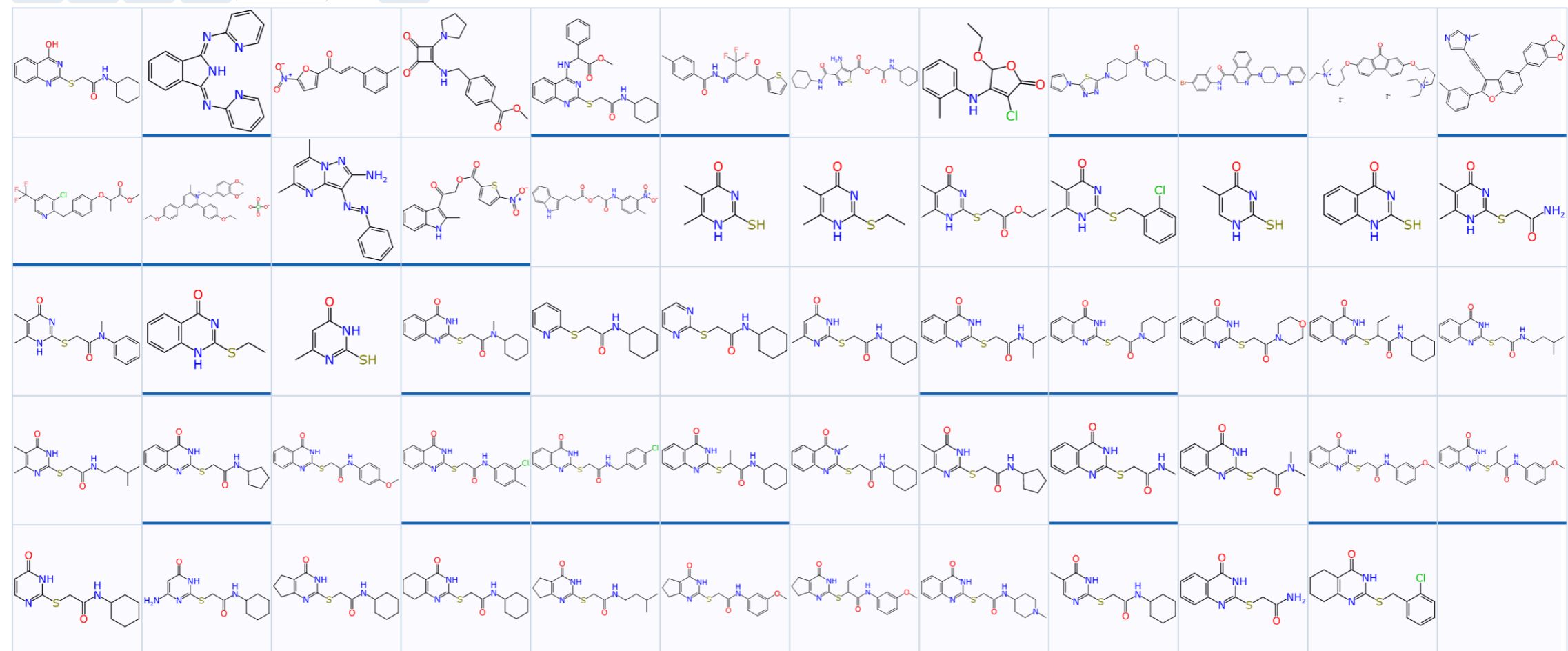
Assay ID: 4011, PubChem Assay ID: 504854, PubChem Source: Southern Research Specialized Biocontainment Screening Center

PubChem Cross References

Taxonomy	https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9534
Other	dburl: http://www.southernresearch.org

+ Annotation History

⏪ ⏩ 1 of 59



Download as SDfile

Now With Compounds

- ◆ Each assay has compounds & data from **PubChem**:

Assay ID: 4011, PubChem Assay ID: 504854, PubChem Source: Southern Research Specia

PubChem Cross References

Taxonomy	https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9534
Other	dburl: http://www.southernresearch.org

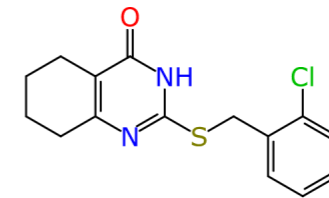
+ Annotation History

1 of 59

The image shows a grid of 59 chemical structures. The 11th structure in the grid is highlighted with a blue border. This structure is a benzothiazine derivative with a chlorine atom on the phenyl ring. The grid is organized into rows and columns, with the highlighted structure being the only one in its row that is highlighted.

Download as SDfile

Compound Measurements



PubChem origins: compound CID 1480154 substance SID 121270649

Name	Type	Relation	Value	Units	
Active/Inactive	activity	=	1	binary	
Average CC50	measurement	=	100	uM	
CC50 Hill Slope Rep 1	measurement	=	-1.05		
CC50 Hill Slope Rep 2	measurement	=	-0.87		
CC50 NormChi2 Rep 1	measurement	=	5.72		
CC50 NormChi2 Rep 2	measurement	=	6.04		
CC50 Rep 1	measurement	=	100	uM	
CC50 Rep 2	measurement	=	100	uM	
CC50 Std Dev Rep 1	measurement	=	107.34		
CC50 Std Dev Rep 2	measurement	=	307.18		
Cell Viability @ 0.195 uM Rep 1	measurement	=	101.34	%	
Cell Viability @ 0.195 uM Rep 2	measurement	=	106.34	%	
Cell Viability @ 0.195 uM Rep 3	measurement	=	99.65	%	
Cell Viability @ 0.391 uM Rep 1	measurement	=	113.19	%	

Now With Compounds

◆ Each assay has compounds & data from **PubChem**:

Measurement Tagging Done

Available Measurements

Assay #1

Measurement	Unit	Activity	Tag
Active/Inactive	binary	activity	<input checked="" type="checkbox"/> IsActive
Average CC50	uM	measurement	<input checked="" type="checkbox"/> CC50_uM
CC50 Hill Slope Rep 1		measurement	<input type="checkbox"/>
CC50 Hill Slope Rep 2		measurement	<input type="checkbox"/>
CC50 Hill Slope Rep 3		measurement	<input type="checkbox"/>
CC50 NormChi2 Rep 1		measurement	<input type="checkbox"/>
CC50 NormChi2 Rep 2		measurement	<input type="checkbox"/>
CC50 NormChi2 Rep 3		measurement	<input type="checkbox"/>
CC50 Rep 1	uM	measurement	<input type="checkbox"/>
CC50 Rep 2	uM	measurement	<input type="checkbox"/>
CC50 Rep 3	uM	measurement	<input type="checkbox"/>
CC50 Std Dev Rep 1		measurement	<input type="checkbox"/>
CC50 Std Dev Rep 2		measurement	<input type="checkbox"/>
CC50 Std Dev Rep 3		measurement	<input type="checkbox"/>
Cell Viability @ 0.195 uM Rep 1	%	measurement	<input type="checkbox"/>
Cell Viability @ 0.195 uM Rep 2	%	measurement	<input type="checkbox"/>
Cell Viability @ 0.195 uM Rep 3	%	measurement	<input type="checkbox"/>
Cell Viability @ 0.391 uM Rep 1	%	measurement	<input type="checkbox"/>

Output Columns

CC50_uM

Average CC50 (measurement, uM)

IsActive

Active/Inactive (activity, binary)

Download as SDfile















Looking at Primary Screens

- ◆ *Browse* the curated content from **PubChem**:

Looking at Primary Screens

◆ Browse the curated content from **PubChem**:











PRIMARY **screening campaign stage**

- 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)























Looking at Primary Screens

◆ Browse the curated content from **PubChem**:

PRIMARY **screening campaign stage**

- 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 **organism**

- absence  
 - unknown (1 of 7)
- organism  
 - cellular organisms  
 - bacterium  
 - Bacillus anthracis (1 of 4)
 - Clostridium botulinum (2 of 21)
 - Escherichia coli (3 of 69)
 - Helicobacter pylori (1 of 10)
 - Leuconostoc mesenteroides (1)
 - Mycobacterium tuberculosis (4 of 54)
 - Pseudomonas aeruginosa PAO1 (3 of 58)
 - Salmonella enterica subsp. enterica serovar Typhimurium (1 of 15)
 - Vibrio harveyi (1 of 25)
 - eukaryote  
 - fungi  
 - Cryptococcus neoformans (1 of 17)
 - multicellular fungi  
 - Candida albicans (1 of 50)
 - unicellular fungi  
 - Saccharomyces cerevisiae (18 of 92)
 - metazoa  
 - invertebrate  
 - Caenorhabditis elegans (2 of 25)
 - Fasciola hepatica (1 of 3)
 - Haemonchus contortus (1 of 3)
 - Photinus pyralis (1 of 17)
 - vertebrate  
 - mammalian  
 - Bos taurus (3 of 7)
 - Canis lupus familiaris (1 of 11)
 - **Homo sapiens** (255 of 2295)
 - Mus musculus (18 of 285)
 - Rattus norvegicus (7 of 89)
 - non-mammalian  
 - Danio rerio (1 of 5)

Matching Assays

Matching assays: 255

All

None

Show Compounds

Assay Grid

Copy Results

Query

Compounds	Result	PubChem
<input type="checkbox"/> 16000	Primary Cell Based High Throughput Screening Assay for Agonists of GALR2	View AID833
<input type="checkbox"/> 150839	qHTS Assay for Inhibitors of HPGD (15-Hydroxyprostaglandin Dehydrogenase)	View AID894
<input type="checkbox"/> 211511	qHTS Assay for Enhancers of SMN2 Splice Variant Expression	View AID1458
<input type="checkbox"/> 194920	Multiplexed high-throughput screen for small molecule regulators of Bcl-2 family protein interactions, specifically Bim-Bcl-XL.	View AID1007
<input type="checkbox"/> 194920	Multiplexed high-throughput screen for small molecule regulators of Bcl-2 family protein interactions, specifically Bim-Bcl-W.	View AID952
<input type="checkbox"/> 194920	Multiplexed high-throughput screen for small molecule regulators of Bcl-2 family protein interactions, specifically Bim-Bcl-B.	View AID951
<input type="checkbox"/> 64925	Primary biochemical high-throughput screening assay for inhibitors of Matrix Metalloproteinase 13 (MMP13) activity	View AID570
<input type="checkbox"/> 302517	MLPCN Platelet Activation -Dense Granule Release	View AID1663
<input type="checkbox"/> 1352	qHTS Assay for Inhibitors of CDC-like Kinase 4 (ADP-Glo Assay)	View AID1771
<input type="checkbox"/> 1352	qHTS Assay for Inhibitors of CDC-like Kinase 4 (Kinase-Glo Assay)	View AID1770
<input type="checkbox"/> 362388	uHTS identification of CXCR6 Inhibitors in a B-arrestin luminescence assay	View AID602244
<input type="checkbox"/> 8590	Inhibitors of USP1/UAF1: Pilot qHTS	View AID504865
<input type="checkbox"/> 345034	A quantitative high throughput screen for small molecules that induce DNA re-replication in MCF 10a normal breast cells.	View AID624296
<input type="checkbox"/> 344074	A quantitative high throughput screen for small molecules that induce DNA re-replication in SW480 colon adenocarcinoma cells.	View AID624297
<input type="checkbox"/> 9984	VCAM-1 Imaging Assay in Pooled HUVECs: Inhibition of TNF α induced VCAM-1 cell surface expression	View AID456
<input type="checkbox"/> 9984	VCAM-1 Imaging Assay in Pooled HUVECs: Augmentation of TNF α induced VCAM-1 cell surface expression.	View AID457
<input type="checkbox"/> 303344	MLPCN Ras selective lethality-BJeLR viability	View AID1554
<input type="checkbox"/> 104742	Inhibitors of Cav3 T-type Calcium Channels: Primary Screen	View AID449739
<input type="checkbox"/> 20540	HTS colorimetric detection of phosphate released in TNAP reaction	View AID614
<input type="checkbox"/> 20540	HTS colorimetric detection of p-nitrophenol released in TNAP reaction	View AID615
<input type="checkbox"/> 64394	HTS discovery of chemical inhibitors of anti-apoptotic protein Bfl-1	View AID432
<input type="checkbox"/> 282587	qHTS for Inhibitors of the Interaction of Thyroid Hormone Receptor and Steroid Receptor Coregulator 2	View AID1469
<input type="checkbox"/> 96416	High Throughput Fluorescence Polarization Screen for Bcl-B Phenotype Converters	View AID748
<input type="checkbox"/> 12344	TNF α Induced E-Selectin Expression - Primary screen	View AID487
<input type="checkbox"/> 218117	Primary cell-based high-throughput screening assay to identify agonists of the transient receptor potential channel ML3 (TRPML3)	View AID1448
<input type="checkbox"/> 221370	qHTS Assay for Antagonists of the Neuropeptide S Receptor: cAMP Signal Transduction	View AID1461
<input type="checkbox"/> 16000	Luminescence-based primary cell-based high throughput screening assay to identify inhibitors of NADPH oxidase 1 (Nox1): Maybridge Library	View AID1792
<input type="checkbox"/> 315101	Fluorescence-based primary biochemical high throughput screening assay to identify inhibitors of Protein Phosphatase 5 (PP5).	View AID1987
<input type="checkbox"/> 291075	uHTS HTRF assay for identification of inhibitors of SUMOylation	View AID2006
<input type="checkbox"/> 291075	uHTS fluorescence polarization assay for the identification of translation initiation inhibitors (eIF4H)	View AID2012
<input type="checkbox"/> 315100	Fluorescence polarization-based primary biochemical high throughput screening assay to identify inhibitors of myeloid cell leukemia sequence 1 (MCL1) interactions with BIM BH2 peptide	View AID2057

Matching Assays

Matching assays: 255

All

None

Show Compound

Assay Grid

Copy Results

Query

Compounds	Result	PubChem
<input type="checkbox"/> 16000	Primary Cell Based High Throughput Screening Assay for Agonists of GALR2	View AID833
<input type="checkbox"/> 150839	qHTS Assay for Inhibitors of HPGD (15-Hydroxyprostaglandin Dehydrogenase)	View AID894
<input type="checkbox"/> 211511	qHTS Assay for Enhancers of SMN2 Splice Variant Expression	View AID1458
<input type="checkbox"/> 194920	Multiplexed high-throughput screen for small molecule regulators of Bcl-2 family protein interactions, specifically Bim-Bcl-XL.	View AID1007
<input type="checkbox"/> 194920	Multiplexed high-throughput screen for small molecule regulators of Bcl-2 family protein interactions, specifically Bim-Bcl-W.	View AID952
<input type="checkbox"/> 194920	Multiplexed high-throughput screen for small molecule regulators of Bcl-2 family protein interactions, specifically Bim-Bcl-B.	View AID951
<input type="checkbox"/> 64925	Primary biochemical high-throughput screening assay for inhibitors of Matrix Metalloproteinase 13 (MMP13) activity	View AID570
<input type="checkbox"/> 302517	MLPCN Platelet Activation -Dense Granule Release	View AID1663
<input type="checkbox"/> 1352	qHTS Assay for Inhibitors of CDC-like Kinase 4 (ADP-Glo Assay)	View AID1771
<input type="checkbox"/> 1352	qHTS Assay for Inhibitors of CDC-like Kinase 4 (Kinase-Glo Assay)	View AID1770
<input type="checkbox"/> 362388	uHTS identification of CXCR6 Inhibitors in a B-arrestin luminescence assay	View AID602244
<input type="checkbox"/> 8590	Inhibitors of USP1/UAF1: Pilot qHTS	View AID504865
<input type="checkbox"/> 345034	A quantitative high throughput screen for small molecules that induce DNA re-replication in MCF 10a normal breast cells.	View AID624296
<input type="checkbox"/> 344074	A quantitative high throughput screen for small molecules that induce DNA re-replication in SW480 colon adenocarcinoma cells.	View AID624297
<input type="checkbox"/> 9984	VCAM-1 Imaging Assay in Pooled HUVECs: Inhibition of TNF α induced VCAM-1 cell surface expression	View AID456
<input type="checkbox"/> 9984	VCAM-1 Imaging Assay in Pooled HUVECs: Augmentation of TNF α induced VCAM-1 cell surface expression.	View AID457
<input type="checkbox"/> 303344	MLPCN Ras selective lethality-BJeLR viability	View AID1554
<input type="checkbox"/> 104742	Inhibitors of Cav3 T-type Calcium Channels: Primary Screen	View AID449739
<input type="checkbox"/> 20540	HTS colorimetric detection of phosphate released in TNAP reaction	View AID614
<input type="checkbox"/> 20540	HTS colorimetric detection of p-nitrophenol released in TNAP reaction	View AID615
<input type="checkbox"/> 64394	HTS discovery of chemical inhibitors of anti-apoptotic protein Bfl-1	View AID432
<input type="checkbox"/> 282587	qHTS for Inhibitors of the Interaction of Thyroid Hormone Receptor and Steroid Receptor Coregulator 2	View AID1469
<input type="checkbox"/> 96416	High Throughput Fluorescence Polarization Screen for Bcl-B Phenotype Converters	View AID748
<input type="checkbox"/> 12344	TNF α Induced E-Selectin Expression - Primary screen	View AID487
<input type="checkbox"/> 218117	Primary cell-based high-throughput screening assay to identify agonists of the transient receptor potential channel ML3 (TRPML3)	View AID1448
<input type="checkbox"/> 221370	qHTS Assay for Antagonists of the Neuropeptide S Receptor: cAMP Signal Transduction	View AID1461
<input type="checkbox"/> 16000	Luminescence-based primary cell-based high throughput screening assay to identify inhibitors of NADPH oxidase 1 (Nox1): Maybridge Library	View AID1792
<input type="checkbox"/> 315101	Fluorescence-based primary biochemical high throughput screening assay to identify inhibitors of Protein Phosphatase 5 (PP5).	View AID1987
<input type="checkbox"/> 291075	uHTS HTRF assay for identification of inhibitors of SUMOylation	View AID2006
<input type="checkbox"/> 291075	uHTS fluorescence polarization assay for the identification of translation initiation inhibitors (eIF4H)	View AID2012
<input type="checkbox"/> 315100	Fluorescence polarization-based primary biochemical high throughput screening assay to identify inhibitors of myeloid cell leukemia sequence 1 (MCL1) interactions with BIM BH2 peptide	View AID2057

Frequent Hitters

- ◆ Rank by **score** = # hits vs. selected assays

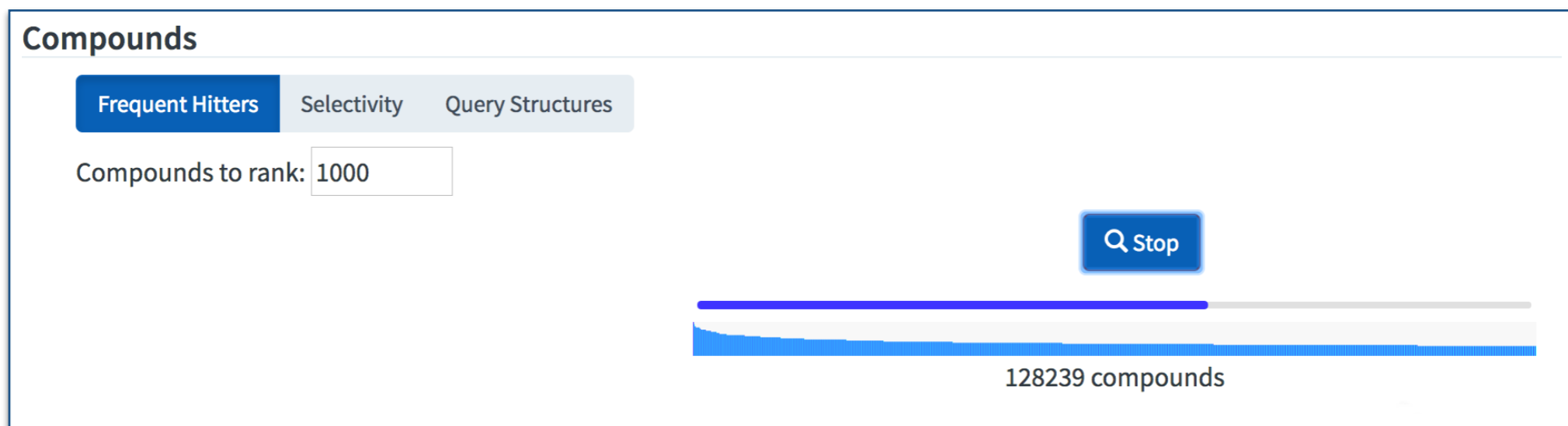
Compounds

Frequent Hitters Selectivity Query Structures

Compounds to rank:

Q Stop

128239 compounds

The image shows a screenshot of a web application interface. At the top, there's a header 'Compounds'. Below it, there are three tabs: 'Frequent Hitters' (which is active and highlighted in blue), 'Selectivity', and 'Query Structures'. Under the 'Frequent Hitters' tab, there is a text input field labeled 'Compounds to rank:' containing the number '1000'. To the right of this field is a blue button with a magnifying glass icon and the text 'Stop'. Below these elements is a horizontal progress bar. The bar is mostly grey, but the left portion is filled with blue. Below the progress bar, the text '128239 compounds' is displayed.

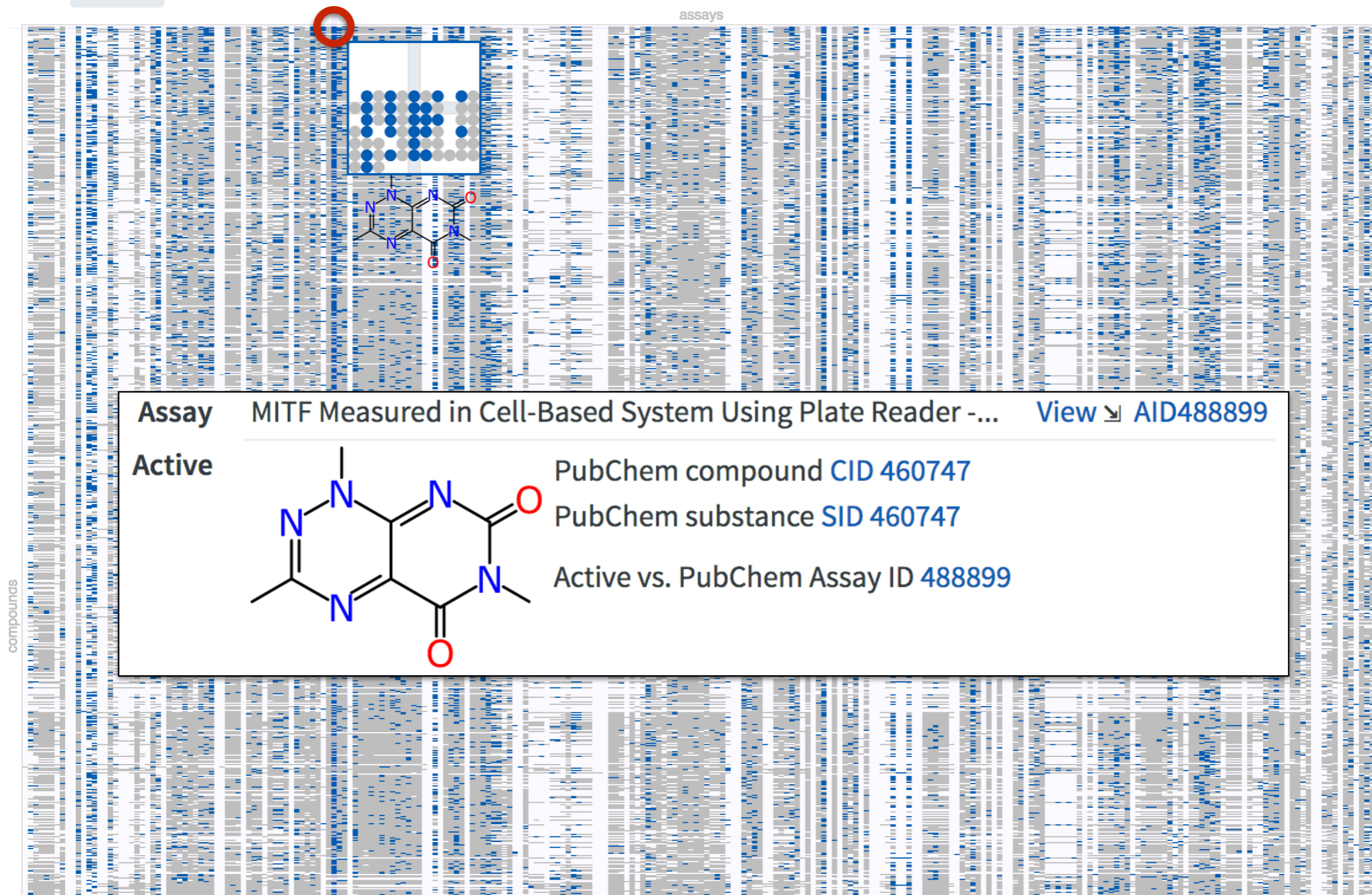
Assay Grid (1°)

Assays: Group By

compounds	assays
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1
50	1
51	1
52	1
53	1
54	1
55	1
56	1
57	1
58	1
59	1
60	1
61	1
62	1
63	1
64	1
65	1
66	1
67	1
68	1
69	1
70	1
71	1
72	1
73	1
74	1
75	1
76	1
77	1
78	1
79	1
80	1
81	1
82	1
83	1
84	1
85	1
86	1
87	1
88	1
89	1
90	1
91	1
92	1
93	1
94	1
95	1
96	1
97	1
98	1
99	1
100	1

Assay Grid (1°)

Assays: Group By



Assay Clustering

- ◆ Leverage hierarchical structure of ontology terms:

Assays:

Assay Clustering

- ◆ Leverage hierarchical structure of ontology terms:

Assays:

Group By

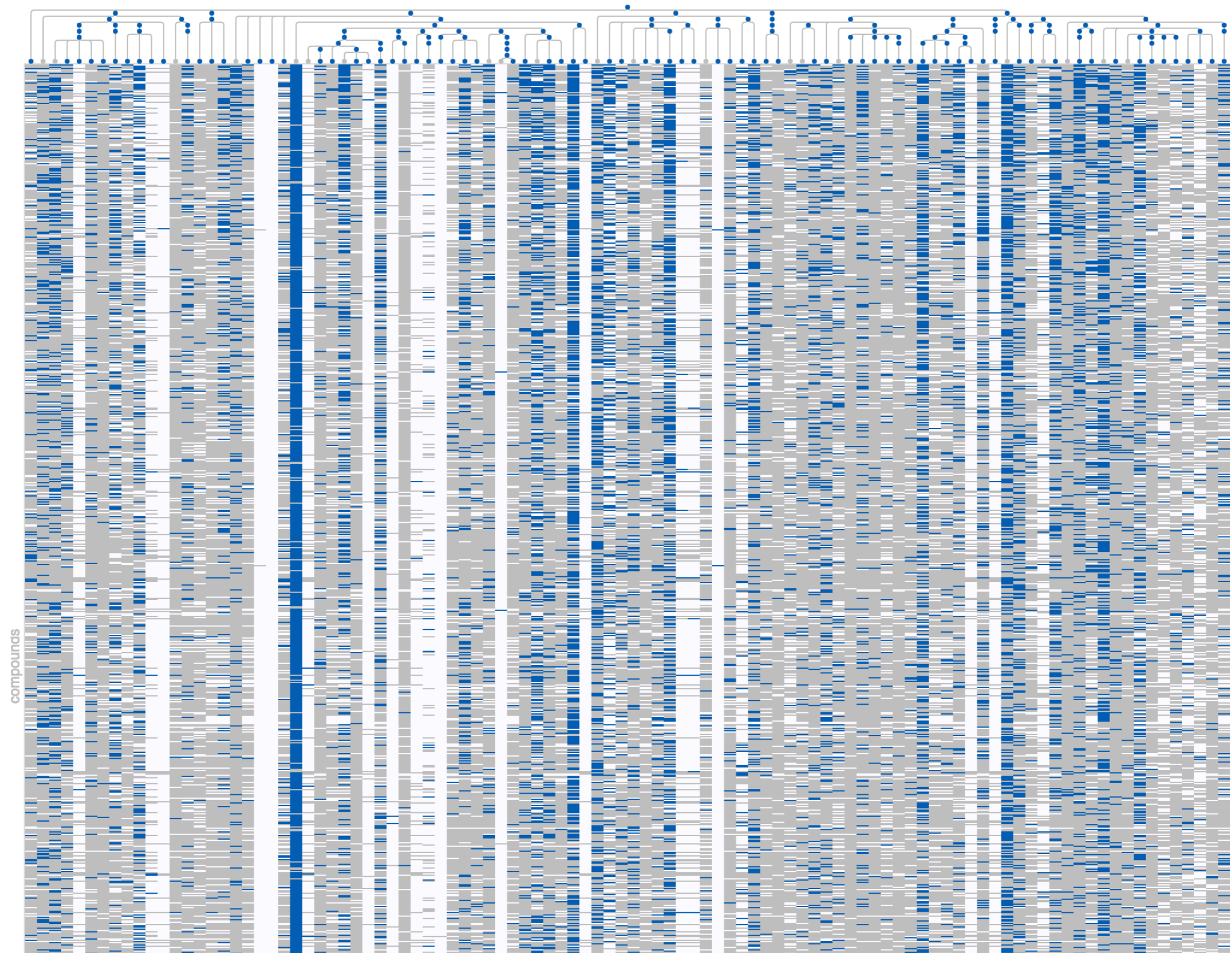


Assay Grid

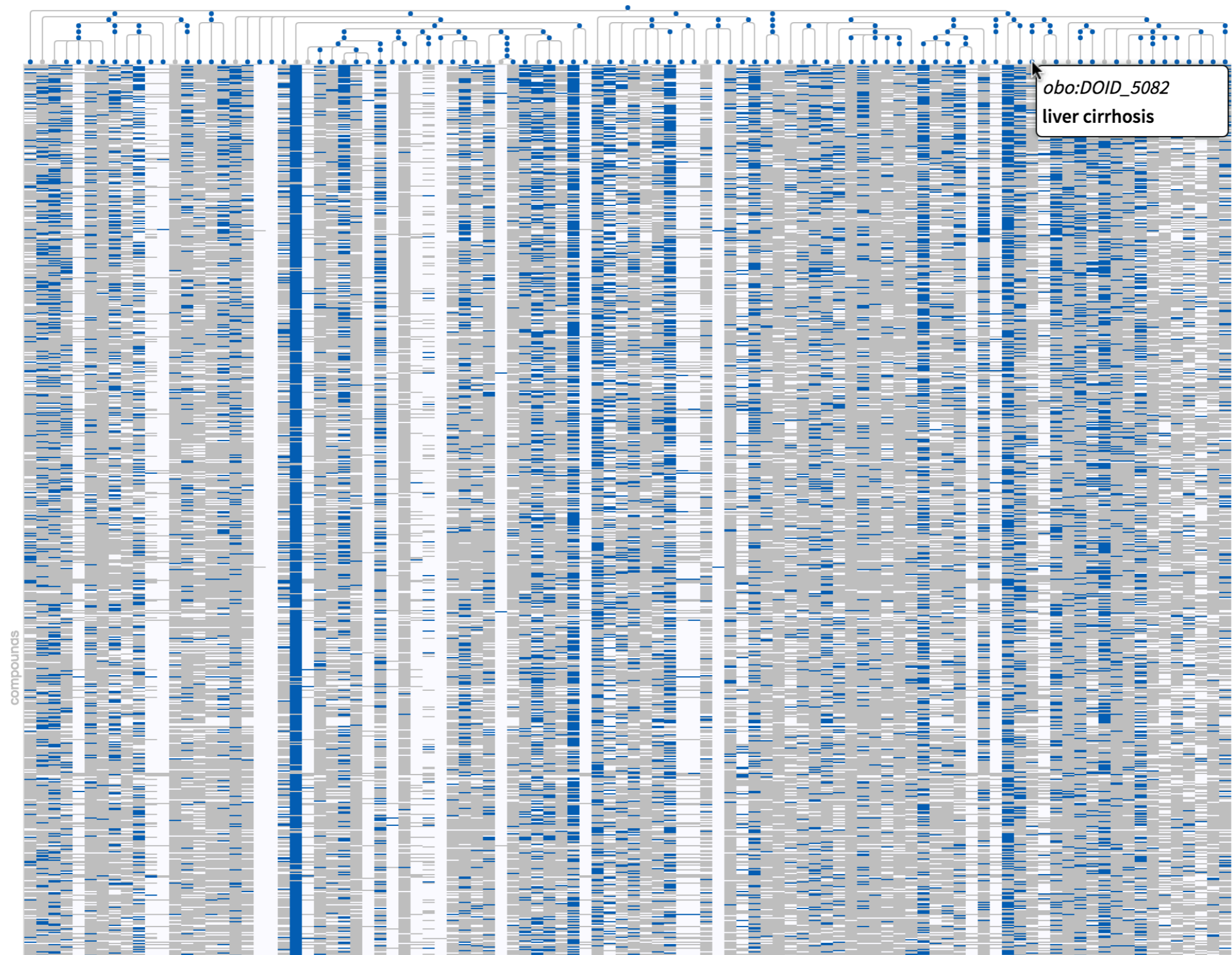
Pick Assignments Done

- bioassay type
- assay format
- assay design method
- assay supporting method
- assay cell line
- organism
- biological process
- target
- applies to disease
- assay mode of action
- result
- measurement unit
- screening campaign stage
- assay footprint
- assay kit
- physical detection method
- detection instrument
- perturbagen type
- protein identity
- gene identity
- GO terms

Group by Disease



Group by Disease



Group by Disease



Selectivity


◆ Using secondary assays

◆ Ranking score =

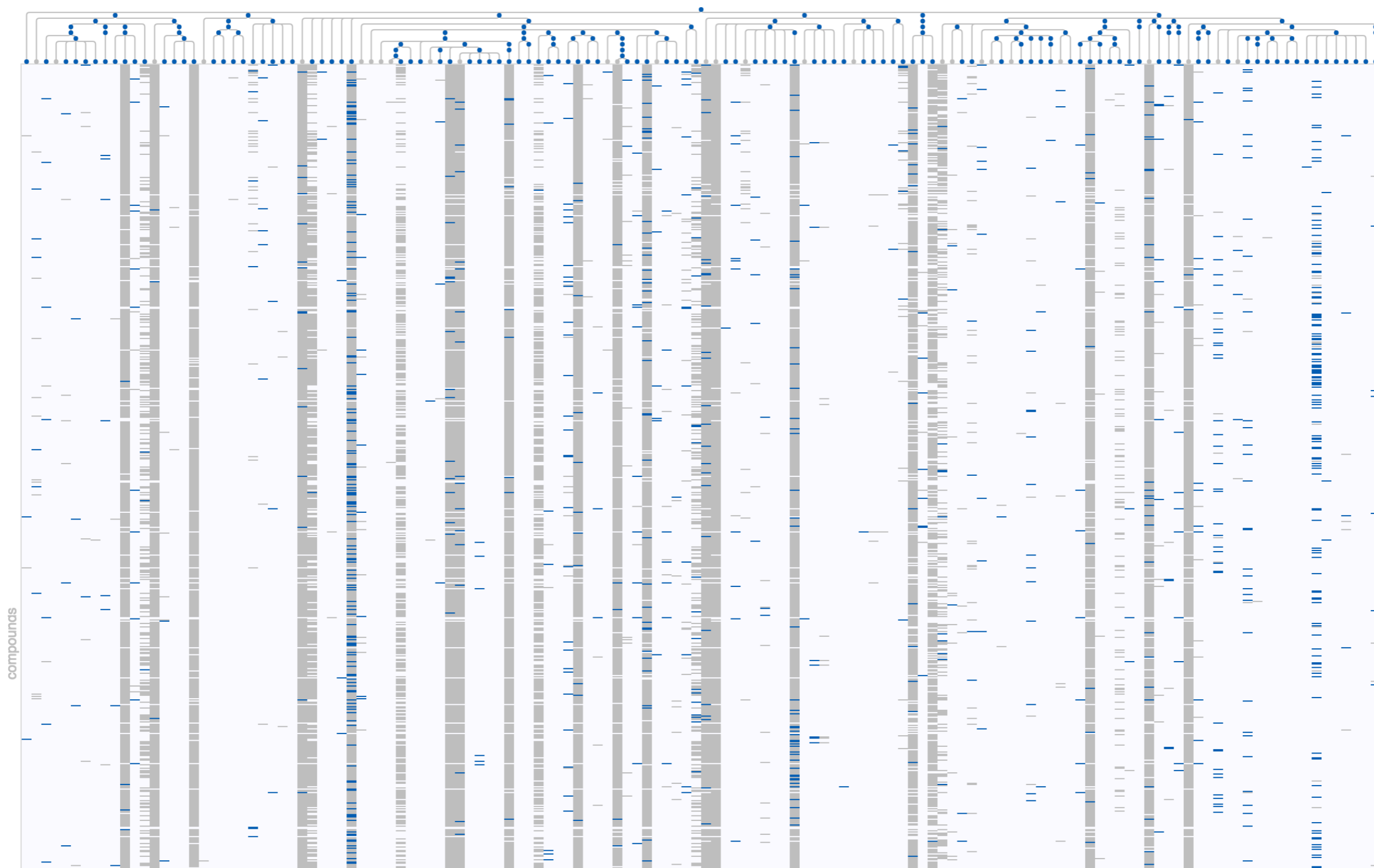
inactives

actives

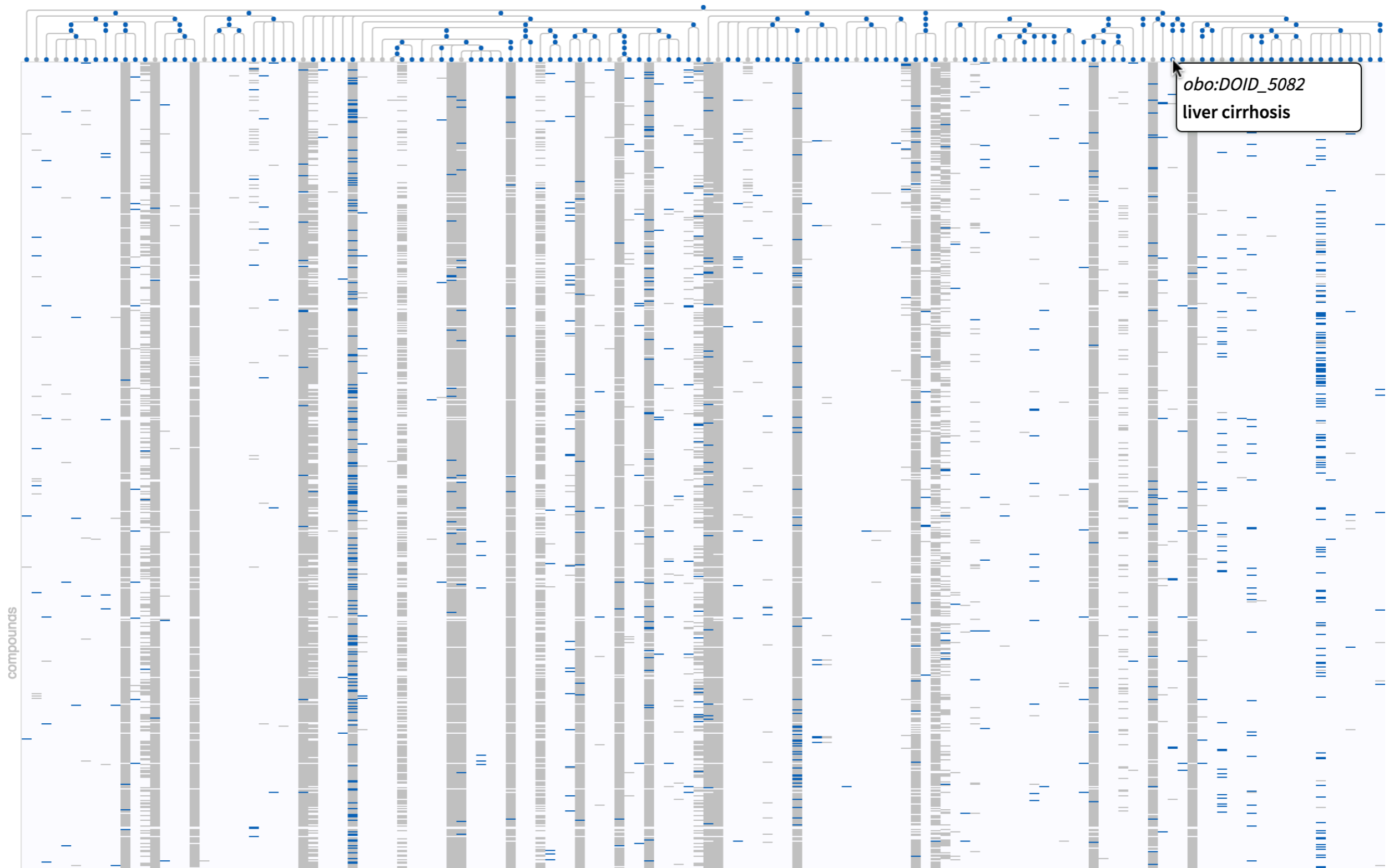
PRIMARY **screening campaign stage**

- 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)

Assay Grid (2°)

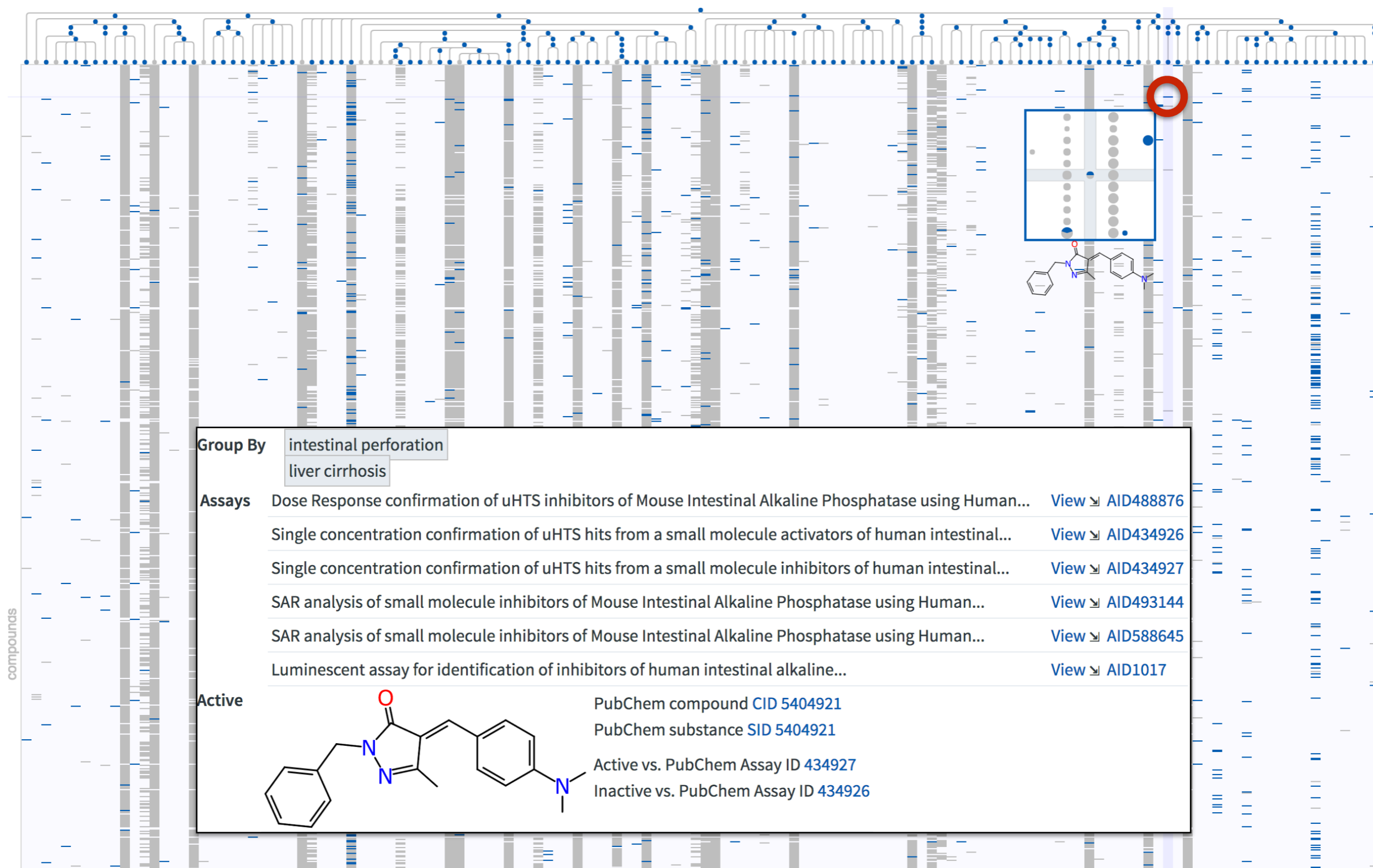


Assay Grid (2°)



obo:DOID_5082
liver cirrhosis

Assay Grid (2°)



Similarity

- ◆ Can select compounds by similarity to a reference structure (ECFP6/Tanimoto):

Assays

Query: (bao:BAX_0000002=@bao:BAO_0002989,@bao:BAO_0000590,@bao:BAO_0003011,@k

Search

Assays: 3510

Compounds

Frequent Hitters

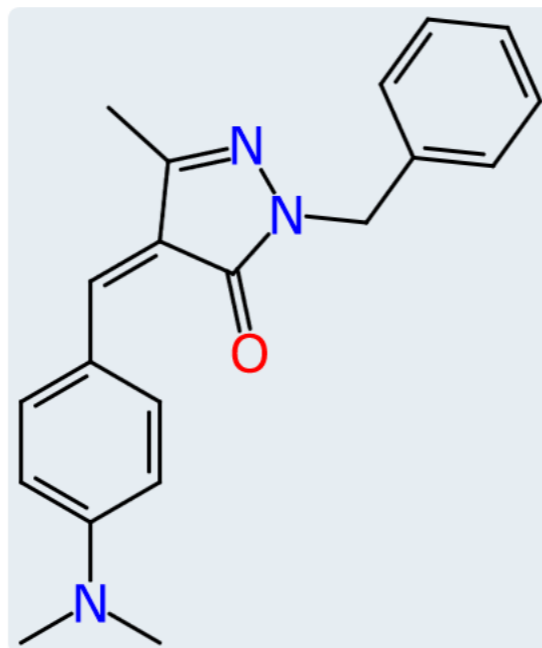
Selectivity

Query Structures

Maximum # of compounds: 1000

Only select probe compounds

Rank by similarity to:



Similarity

- ◆ Can select compounds by similarity to a reference structure (ECFP6/Tanimoto):

Assays

Query: (bao:BAX_0000002=@bao:BAO_0002989,@bao:BAO_0000590,@bao:BAO_0003011,@k

Search

Assays: 3510

Compounds

Frequent Hitters

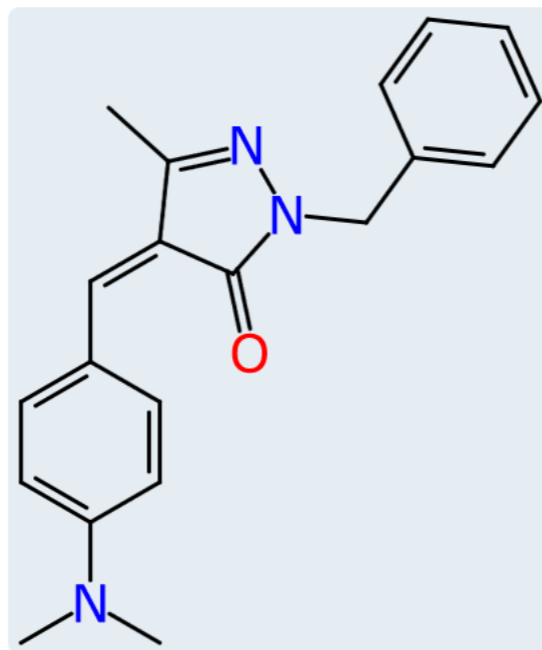
Selectivity

Query Structures

Maximum # of compounds: 1000

Only select probe compounds

Rank by similarity to:



← paste/drag structure

Assay Grid

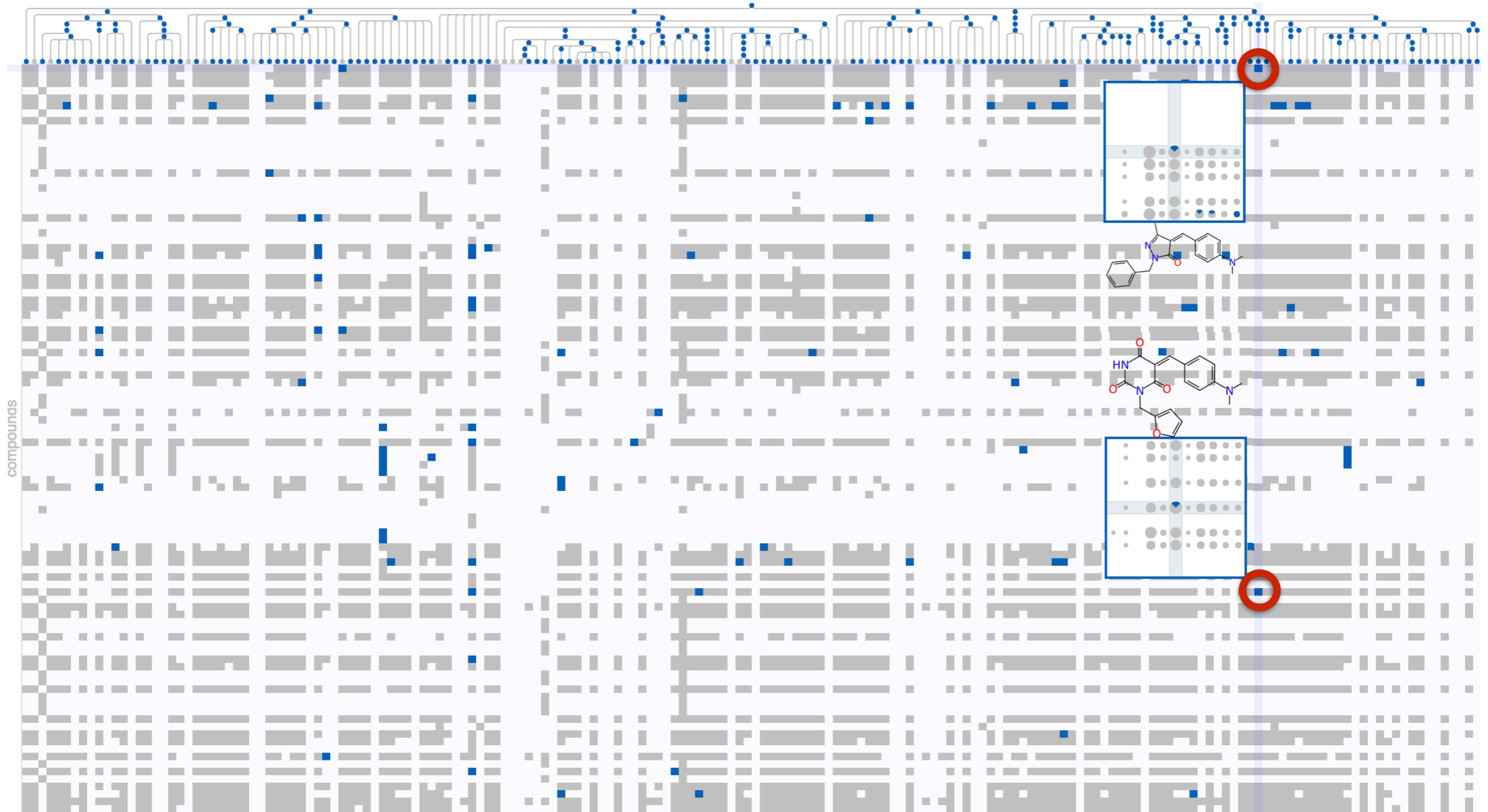
liver
cirrhosis



Group By
intestinal perforation
liver cirrhosis

Assay Grid

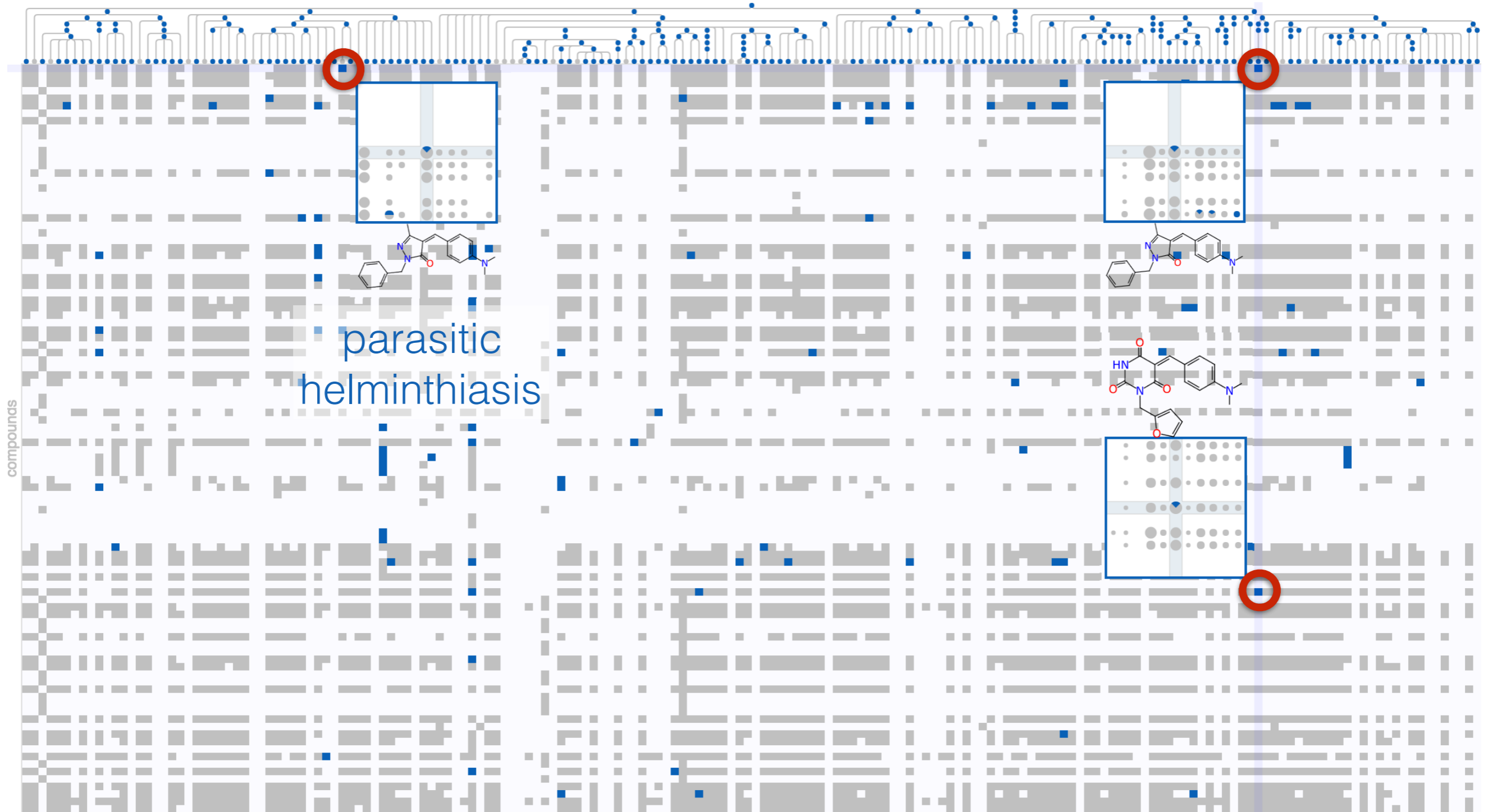
liver
cirrhosis



Group By
intestinal perforation
liver cirrhosis

Assay Grid

liver
cirrhosis



Group By
intestinal perforation
liver cirrhosis

Conclusion

- ◆ *Cheminformatics & bioinformatics* are well established, but **assay informatics** is new
- ◆ Data curation has begun: much yet remains (>1M)
- ◆ Content **curation** is the main bottleneck
 - ▷ *BioAssay Express* is key to solving this problem
- ◆ Content **consumption** is at proof of concept stage
 - ▷ ... and this is the fun part.

Future Work

- ◆ **Crowd curation** of public data: you should want **your** assays to be machine readable
- ◆ **Model** building: create from structure-activity-assay, apply all existing knowledge to prospective discovery
- ◆ **ELN** context: use semantic annotations *instead of* text, describe assays *comprehensively*
 - ▷ convert annotations to text, for communication
 - ▷ translate protocols to import formats for robots

Assay Data

- ◆ All features, and curated data, available at:
 - <http://www.bioassayexpress.com>
 - <http://beta.bioassayexpress.com>
- ◆ Ready for crowd curation - free to use in public
- ◆ Commercial product: private installations, with custom settings, are being deployed

Acknowledgments

◆ Biologists

- ▷ Janice Kranz
- ▷ Haifa Ghandour
- ▷ Karen Featherstone

◆ Collaborative Drug Discovery

- ⇒ Barry Bunin
- ⇒ Hande Küçük
- ⇒ and the rest of the team

◆ More information

<http://www.bioassayexpress.com>

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

<http://collaborativedrug.com>

*PeerJ
Comp Sci
2:e61
(2016)*