Application of an *in silico* Mechanism-of-Action Protocol to High-Content Cytotoxicity Screening Data utilizing WikiPathway Data Extracted from Open PHACTS

Lewis H. Mervin[1], Qing Cao[2], Ian P. Barrett[3], Mike A. Firth[3], David Murray[4], Lisa McWilliams[4], Ola Engkvist[2] and Andreas Bender[1]

[1]Centre for Molecular Informatics, University of Cambridge, UK
[2]Discovery Sciences, Chemistry Innovation Centre, AstraZeneca R&D, Sweden
[3]Discovery Sciences, AstraZeneca R&D, Cambridge Science Park, UK
[4]Discovery Sciences, Innovative Medicines, Alderley Park, Macclesfield, UK

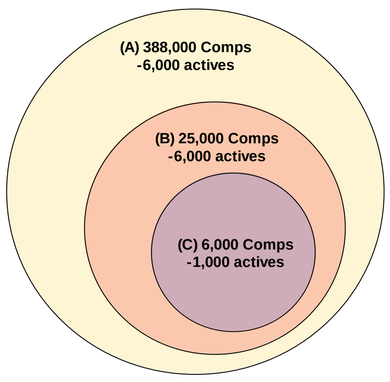
Background

Phenotypic cell-based screening experiments require subsequent target deconvolution and mechanism-of-action (MoA) studies to identify cellular targets and pathways lending to the observed biological response. In this study, we developed an *in silico* MoA analysis protocol, comprising Bernoulli Naïve Bayes compound bioactivity profiling utilising over 9.5 million active and 600 million inactive data points, annotation of predicted targets with WikiPathways extracted from Open PHACTS v2.0, and calculation of enrichment metrics to highlight targets and pathways likely to be implicated in the cytotoxic phenotype. A cytotoxicity prediction model has also been built that will be used within AZ to classify novel compounds.

Method

Active data	Confirmed Inactive data		Putative Inactive data
In-house actives • **9,570,000** actives • **2,882** Targets	**In-house inactives** • **189,500,000** inactives • **343** Targets	**PubChem inactives** • **420,000,000** inactives • **2,116** Targets	**PubChem putative inactives** • **11,000,000** inactives • **800** Targets

Cytotoxicity Data



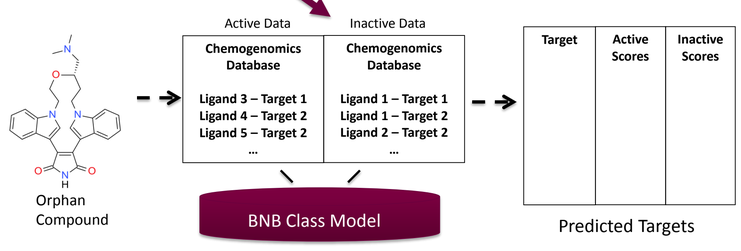
(**A**) Alamar Blue HTS Primary Cytotoxicity Dataset (A)

(**B**) Alamar Blue Dose-Response Cytotoxicity Dataset

(**C**) Cell Tox Green Membrane Integrity Assay (C)

Non-toxic dataset = (**A**) inactives
Toxic dataset = (**B**) actives – (**C**) actives

1.) Predict Targets

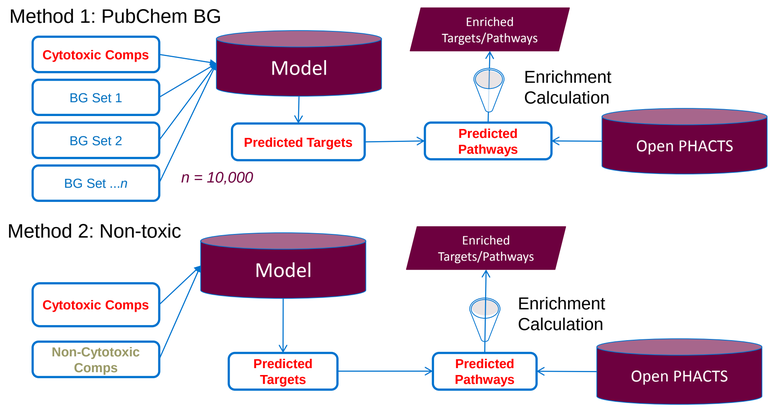


https://github.com/lhm30/PIDGIN [1]

Enrichment Calculation

$$Odds\ Ratio = \frac{F_t / N_t}{F_b / N_b} \qquad Average\ Ratio = \frac{1}{n}\sum_{i=1}^{n}\frac{R_i}{F_t}$$

2.) Calculate Enrichment



Results & Discussion

ID	Name	Hits	Average Ratio
WP2765	Mitotic Telophase	1420	0.037
WP1874	Nucleosome assembly	1921	0.045
WP1861	mRNA Capping	3910	0.046
WP2673	Interleukin-7 signaling	3238	0.051
WP3348	Macroautophagy	17387	0.051
WP438	Non-homologous end joining	1184	0.053
WP405	Eukaryotic Transcription Initiation	3965	0.054
WP1906	RNA Polymerase II Transcription	7116	0.059
WP2760	Signaling by BMP	5600	0.060
WP2748	Energy dependent regulation of mTOR by LKB1-AMPK	16641	0.060
WP1471	TOR Signaling	23108	0.061
WP2732	Interleukin-2 signaling	8846	0.061
WP1845	MAPK targets	3370	0.063
WP1980	Nucleotide Excision Repair	4478	0.064
WP2801	MyD88 cascade initiated on plasma membrane	8713	0.065
WP1808	DSCAM interactions	9011	0.065
WP2276	Glial Cell Differentiation	690	0.070

ID	Name	Hits	Neg. Hits	Odds Ratio
WP3376	RHO GTPases activate CIT	1271	4736	0.081
WP1425	BMP Signalling and Regulation	3240	14524	0.098
WP3352	Toll Like Receptor 3 (TLR3) Cascade	377	1963	0.114
WP3390	Uptake and function of anthrax toxins	4646	25328	0.119
WP2741	Inositol phosphate metabolism	133	745	0.122
WP405	Eukaryotic Transcription Initiation	5013	33215	0.145
WP2798	Assembly of collagen and multimeric structures	122	813	0.145
WP3383	Regulated Necrosis	1574	11029	0.153
WP453	Inflammatory Response Pathway	5278	38268	0.158
WP2007	Iron metabolism in placenta	75	547	0.159
WP3358	Ligand-dependent caspase activation	2422	17727	0.160
WP2760	Signaling by BMP	7147	52625	0.161
WP2113	Type III interferon signaling	4155	30842	0.162
WP1991	SRF in Smooth Muscle Differentiation /Proliferation	2377	18132	0.166
WP2673	Interleukin-7 signaling	3871	29768	0.168

Enriched Pathway Results from the MoA Protocol

(Top Left) Method 1: The highest enriched pathways have biological links to the phenotype at different absolute frequencies. Pathways such as mitotic telophase[2], interleukin signalling[3], macroautphagy, RNA transcription are highlighted. These have obvious links to the cytotoxic MoA, which can be supported by literature review

(Bottom Left) Method 2: The top enriched pathways for the non-toxic background highlight more specific processes. Pathways such as regulated necrosis[4] transcription initiation, interferon signalling, and assembly of collagen, are highlighted, with known links to literature.

Future Work

• Consider methods to improve the pathway overlaps beyond relatively simple target/canonical pathway overlaps

• Consider weighting of pathway members based on other information, such as targets directionality of involvement in the pathway of interest (e.g. an inhibitor of an enzyme which positively relates to a pathway, in turn relates positively to the phenotype may get a better weighting when looking at an assay focussed on reduction of the phenotype)

• Address more potentially causally (inferred) related signals rise above various associative information.

•Cytotoxicity prediction will be used within AZ to classify novel compounds.

References

[**1**] **Mervin**, **Lewis H.** *et al.* *"Target Prediction Utilising Negative Bioactivity Data Covering Large Chemical Space."* Journal of Cheminformatics, 2015, 7:51
[**2**] Warrener, Robyn, et al. "*Tumor cell-selective cytotoxicity by targeting cell cycle checkpoints.*" The FASEB journal, 2003, 17.11
[**3**] Farrar, William L. *et al.* "*Regulation of the production of immune interferon and cytotoxic T lymphocytes by interleukin 2.*", 1981, Journal of Immunology 126.3
[**4**] Nikoletopoulou, Vassiliki, *et al.* "*Crosstalk between apoptosis, necrosis and autophagy.*" Molecular Cell Research, 2013, 1833:12

Acknowledgements

BBSRC and AstraZeneca are thanks for funding. The Bender group and the AstraZeneca team are thanked for their support