Chemical probe assessments for expanding the drug target envelope

Christopher Southan, Deanery of Biomedical Sciences, University of Edinburgh, UK, Current address, Medicines Discovery Catapult, Macclesfield, UK, **Ctibor Škuta** and **Petr Bartůněk**, CZ-OPENSCREEN: National Infrastructure for Chemical Biology, Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic. Abstract ID: 1181, Presentation P017.

Introduction. From 2005-14 the NIH Molecular Libraries Program identified the first generation of small-molecule chemical probes to facilitate the exploration of new targets, and pathways. These successes inspired other organisations to initiate similar efforts from ~2010. However, a decade later it has become difficult to get an overview of exactly what probes are available from where and what proportion of the human druggable proteome they cover.

Methods. The Probes & Drugs portal (P&D), downloaded, integrated and compared sets of declared probe compounds harvested from 12 different sources [1]. We developed analytical approaches to address key questions including; a) individual and total compound structure counts, b) overlaps between sources, c) comparisons with selected PubChem sources and d) investigating the probe coverage of druggable targets and d) new comparative quality scores.

Results. Our analysis of 944 experimental and 3,670 calculated probe candidates provides evidence of specific binding for 796 human proteins across the target classes [2]. We have flagged unsuitable (i.e. potentially misleading and resource-wasting) compounds from both probe groups. Comparison with PubChem revealed unexpectedly high matches with compounds extracted from patents. In addition we developed new high-level scoring schemes to filter collections down to probes of higher quality. This allowed us to generate 550 High-quality chemical probes covering 447 distinct protein targets, now listed on the P&D website. The source overlaps are shown below.

	Set	1	2	3	4	5	6	7	8	9	10	11	12
1	Bromodomains toolbox	25	16	0	0	0	0	1	0	0	21	10	0
2	Chemical Probes.org	16	362	17	2	13	24	10	25	13	43	114	0
3	Gray Laboratory	0	17	53	0	2	0	0	1	0	0	7	0
4	MLP	0	2	0	375	3	0	0	4	0	0	4	4
5	Nature Chemical Biology	0	13	2	3	58	1	0	0	1	4	9	1
6	Open Science Probes	0	24	0	0	1	83	12	2	0	5	0	0
7	opnMe Portal	1	10	0	0	0	12	55	1	0	3	2	0
8	Probe Miner	0	25	1	4	0	2	1	3187	1	2	32	9
9	Methyltransferases toolbox	0	13	0	0	1	0	0	1	19	19	11	0
10	SGC Probes	21	43	0	0	4	5	3	2	19	81	26	2
11	Tool Compound Set	10	114	7	4	9	0	2	32	11	26	515	1
12	Historical Compounds	0	0	0	4	1	0	0	9	0	2	1	239

A matrix showing the intersections between 12 sources using the InChIKey exact match for the standardised structures from the P&D portal. The diagonal figures in white represent the source counts.

Conclusions. Probes offer increasing opportunities to elucidate the functions of understudied proteins which in turn can lead to the validation of new drug targets. The analysis presented here and the updated integration of probe sources will assist researchers in the best choices of probe compounds and their orthogonal controls for insightful experimentation on new targets.

References. [1] Skuta C, Popr M, Muller T, Jindrich J, Kahle M, Sedlak D, Svozil D, Bartunek P (2017) Probes &Drugs portal: an interactive, open data resource for chemical biology. Nature Methods, 14(8):759-760, PMID: 28753599,

[2] Skuta C, Southan C, Bartunek P (2021). Will the chemical probes please stand up? RSC Medicinal Chemistry (accepted) and ChemRxiv https://chemrxiv.org/engage/chemrxiv/article-details/60c7582f4c89196708ad4b3a