How Many Drug Targets Are There? A Status Update Review

Nov. 22, 2017 by Alfred Ajami

The question is often raised, but the answer remains to be uncovered because the definition of drug "target" continues to evolve. Historical conceptualization is focused on catalytic sites, substrate binding sites, or epigenetic modification sites. Current understanding that protein-protein interactions are druggable, along with the emerging realization that "nodes" in signaling pathways and biological networks themselves can be manipulated with small molecules in non-traditional ways, has opened up new targeting options. This review is intended to provide a status update, and you can also access a list of 36 actionable web resources for target hunting.

The first comprehensive, and high visibility, attempt to quantify the number of targets, based on the pharmacopoeia of approved drugs, appeared in 2006 and has since garnered 1850+ citations, elevating it almost to the status of dogma. Overington and colleagues concluded on the "basis of existing knowledge" that "all current drugs with a known mode-of-action act through 324 distinct molecular drug targets" (Ref. 1). Subsequent efforts along the same lines of reasoning over the last decade have increased the practical target count to the 500-900 range (Refs. 2-3) as a subset of the overall druggable genome and generally accepted putative target count of 2000-3000 genes (Ref. 4), although the latter has now been expanded to ~4500 (Ref. 5).

An alternative census also has emerged in step with the growth of systems biology. Analysis of signaling pathways and networks led Fazekas et al. to conclude in 2013 that the targetable genome and epigenome harbor at least 14,392 non-redundant intervention sites amenable to modulation, a third of which consist of post-translational modifiers (Ref. 6). Earlier this month, Shih et al. published their exhaustive curation and an encyclopedic review of the Cortellis database (Clarivate Analytics) to map the drug-project landscape (Ref. 7). Derek Lowe offered a concise summary in his commentary on the paper: "That boils down to about 2,400 mechanisms for 1,400 indications, and the combination of those two yields about 15,000 mechanism/indication pairs", which can be regarded as another version count of the drug-disease targeting superset (Ref. 8)

In summary, it appears that conservative definitions based on known or approved drug mechanisms (molecular and cellular) yield a target count close 1,000; system/network biology and a more liberal definition of "target" put the count at close to 15,000. The reductionist might split the difference and

default to a curated target count from ChEMBL: 7,610. This is a subset of nominal total count of 11,538 in the current ChEMBL statistical summary, and the breakdown is given in the Table below (Ref. 9).

	Targets in ChEMBL
Adhesion	18
Auxilliary transport protein	26
Enzyme	3983
Epigenetic regulator	155
Ion Channel	441
Membrane receptor	948
Other cytosolic protein	98
Other membrane protein	15
Other nuclear protein	11
Secreted protein	87
Structural protein	32
Surface antigen	29
Transcription factor	195
Transporter	260
Unclassified protein	1312
TOTAL	**7610**

If the trend continues and more targets are both uncovered and validated in step with advances in CRISPRi/a screening, single-cell sequencing, and proteomics, coupled with big data management and artificial intelligence (AI) discovery, the target census should be expected to expand well into five figures.

Please, leave your comments below to start a conversation.

References:

1. How many drug targets are there?, Overington et al., Nature Revs. Drug Disc. 2006 (not O/A), http://go.nature.com/2hMX2Sv; (available on researchgate.net, http://bit.ly/2AVheWd)
2. The druggable genome: Evaluation of drug targets in clinical trials suggests major shifts in molecular class and indication, Rask-Andersen et al., Ann. Rev. Pharmacol. Toxicol. 2014 (not O/A), http://bit.ly/2zYLSP9
3. A comprehensive map of molecular drug targets, Santos et al., Nature Revs. Drug Disc. 2017 (not O/A), http://go.nature.com/2AePzmX (available on semanticscholar.net, http://bit.ly/2jJm5Xn)

4. The druggable genome: an update, Russ and Lampel, Drug Disc. Today 2015 (not O/A), http://bit.ly/2jGwdzY
5. The druggable genome and support for target identification and validation in drug development, Finan et al., Science Trans. Med. 2017 (not O/A), http://bit.ly/2hQ5Ua3 (available in manuscript form on bioRxiv, http://bit.ly/2zlZFlY)
6. SignaLink 2 - a signaling pathway resource with multi-layered regulatory networks, Fazekas et al., BMC Syst. Biol. 2013 (O/A), http://bit.ly/2hNkXBe; portal: http://signalink.org
7. Drug discovery effectiveness from the standpoint of therapeutic mechanisms and indications, Shih et al., Nature Revs. Drug Disc. 2017 (not O/A); http://go.nature.com/2AgsDnf
8. "In the Pipeline" blog commentary, Lowe 2017 (O/A), http://bit.ly/2zZ6l6i
9. The ChEMBL database in 2017, Gaulton et al., Nucleic Acids Res. 2017 (O/A), http://bit.ly/2AhBcxV; portal for target browser: http://bit.ly/2zVIUNy; blog about drug targets: http://bit.ly/2hGgGvN. The graphic below shows the target browser and selection screen from ChEMBL_20. The current version is ChEMBL_23.