The Ontology Reference Model for Visual Selectivity Analysis in Drug-target Interactions

Qiong Cheng[1], Felix A Lopez[2], Celia Duran[3], Christopher Camarillo[4]
School of Engineering and Technology
Miami Dade College
[1]qcheng, [2]flopez@mdc.edu, [3]celia.duran001, [4]chris.camarillo001 @mymdc.net

Tudor I. Oprea[4]
Translational Informatics Division, Department of Internal Medicine
University of New Mexico
Albuquerque, NM 87131
[4]toprea@salud.unm.edu

Stephan Schurer[5]
Department of Molecular and Cellular Pharmacology
University of Miami
Miami, USA
[5]stephan.schurer@gmail.com

Abstract—In a drug development process, appropriate drug-binding selectivity is critical for a success drug. However the selectivity in a data source, showing the intensity of efforts, may be limited to prior knowledge of the expertise or be biased towards the hypothesis testing. With the increasing of drug screening data, it is challenging to coordinate the efforts and execute data governance at a large scale. Visual selectivity analysis for examining target selection is in demand. We proposed a knowledge-driven approach and designed an ontology reference model to provide an intuitive view of the selectivity in a drug-target interaction network data. We employed the model to carry out the visual selectivity analysis on the NIMH Psychoactive Drug Screening Program (PDSP) data and the LINCS Compounds-interacting ChEMBL Database. The analysis indicates the possible 'dark matter' drug targets. The approach can be expanded to coordinate other experimental screening data and set a stage for the analysis of the mechanism of action of biological therapies.

Keywords—drug-target interactions, ontology, binding selectivity

I. INTRODUCTION

A mainstream drug discovery process is an assembly pipeline starting with target identification, following with lead identification and optimization, and completing with preclinical and clinical assessment [17]. In the process, target selection for the screening is a key factor for the successful drug development since it can be used in the early drug development stage for appropriate balance of avoidance of undesirable targets and coverage of one or more targets of interest.

A large and multifaceted effort in biomedical research has emphasized on target selection and further brought forth the prosperity of drug discovery. However, in one hand, among the estimated 20,000-25,000 human protein-coding genes, there are only ~3% being considered as the targets of approved drugs with known mechanism of action [4]. In another hand, there are still ~95% rare diseases and ~50% genetic diseases that have no a single FDA approved drug associated. The main reason possibly lies in the existence of understudied targets or the poor or biased target selection, which causes the efficacy issues. In some instances, the actual target may not be known; a drug may have a primary target identified but may interact with other proteins and molecules. The number of proteins and other molecules with which a drug interacts throughout the cell may be large. A drug-target experimental data may be biasedly towards a subset of proteins being selected to screen against a chemical library, which cause drug toxicity being found in the late drug discovery. The type of drug toxicity is caused by the unwanted off-targets. Namely, the candidate compound binds to and influences not only the desired target but also other unintended proteins causing side effects [18-21]. Identifying and prioritizing on/off targets is challenging and plays a critical role in drug discovery.

Many drug candidates are found to be unsafe or lack of efficacy only late in the drug discovery process, which is the reason that developing a new drug from original idea to the launch of a finished product may take 12–15 years and cost in excess of 1 billion dollars [22]. A knowledge base-driven systematic analysis in earlier phases in evaluating the selection and effects will be beneficial for a better decision in the early drug discovery process.

An intuitive approach to identify drug selection would be to screen the entire human proteome with the drug candidate in question; however, economic and technological requirements would make such an approach unfeasible [17]. Thus, target selection in many experimental drug screening data is biased towards the hypothesis testing or depends on prior knowledge of the expertise, showing the intensity of efforts from drug-target data source.

With the detailed sequence of the Human Genome available, the druggable genome has attracted more and more interest of scholars. The druggable genome is defined as the subsets of genes in the human genome that express proteins capable to bind drug-like small molecules [19]. The Illuminating the Druggable Genome (IDG) project (http://targetcentral.ws/) has

978-1-5090-3050-7/17/$31.00 ©2017 IEEE

developed an integrated knowledge base for the human druggable genome with a goal of improving our understanding of the properties and functions of proteins that are currently unannotated within the four most commonly drug-targeted protein families: the G-protein coupled receptors (GPCRs), nuclear receptors, ion channels, and protein kinases [12, 15-17]. In the context of the IDG program, the Drug Target Ontology (DTO) was developed in [14], which provides a structured knowledge resource of drug target classifications and relevant annotations for these four protein families in the IDG knowledge base. In the paper, drug targets are defined as the druggable proteins that possess structural and functional features of druggability.

Numerous target-based drug profiling datasets have been generated through the efforts of a panel of protein targets against various laboratories and are publically available via different databases such as NIMH Psychoactive Drug Screening Program (PDSP) data [1, 2], Kinome Knowledge Base (KKB) [7], KINOMEscan and ChEMBL [6]. The HMS LINCS Center [8] provided phenotype-based drug perturbation profiling data that signals cellular response to drug perturbation.

It is challenging for a data coordinator center to execute data governance with the diversity in consideration. We propose a knowledge-driven approach to facilitate data governance, in which the availability, usability, integrity, and security of a drug-target data set are examined before available in public. In the paper, we focus on examining the consistency or integrity at diverse levels. We use the druggable genome described in the IDG project as the reference to examine the number of selected targets in a drug target data set and further visualize its coverage in the classification tree of drug targets presented in Drug Target Ontology (DTO); we also employ the statistical method to measure whether compounds interacting with targets within the same protein category of DTO are more similar to those across different categories (manuscript is in review). In a data coordinator center, it is important to validate these uploaded experimental data for better data governance.

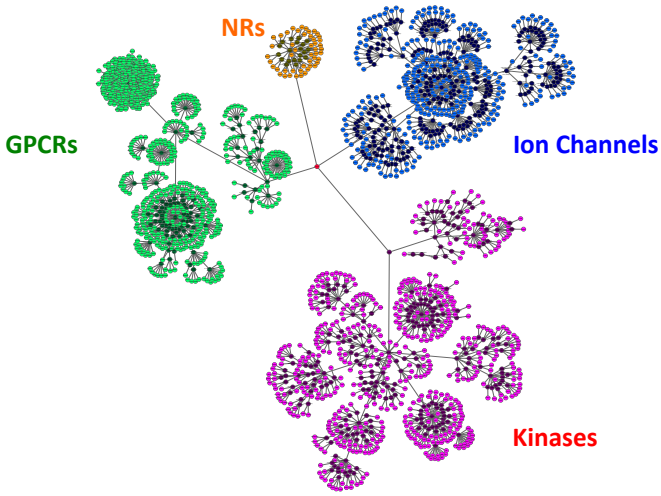

Figure 1. Visualization of druggable targets and its classification tree in Drug-Target Ontology (DTO). Four important gene families, such as G protein-coupled receptors (GPCR), nuclear receptors, ion channels and kinases, are shown in different colors. GPCR proteins and classification tree is in green, Kinases in pink, ion channels in blue, and nuclear receptors in brown.

II. Ontology Reference Visualization Model

A. DTO: Integrative Drug-Target Ontology

We extracted the druggable proteins and its classification tree from the DTO, which is an integrative ontology of coordinating, mapping and integrating protein target concepts across a multitude of relevant data resources and external ontologies with individual curation [14]. In the DTO, the druggable proteins are cast into four important drug target protein families such as G protein-coupled receptors (GPCRs), nuclear receptors, ion channels, and kinases, all of which are commonly targeted for drug development (see Figure 1).

The GPCR classification reflects the phylogenetic, functional, and ligand-type properties through the structure of class, group, family, and subfamily. Most of the information has been taken from the GPCR.org classification and had been updated using various sources *e.g.* IUPHAR [3], ChEMBL, UniProt, and ChEBI. The Nuclear Receptor classification, originally derived from IUPHAR. The Ion Channel classification is primarily structured into family, subfamily, sub-subfamily. It is mostly taken from the Transporter Classification Database, capturing the information of phylogenesis, function, structure, topology, gating mechanism and transported ions. The Kinase classification is formed into several groups, families, and subfamilies with the subsumption hierarchy, being extracted and curated from UniProt, ChEMBL, PhosphoSitePlus® (PSP) [5], Sugen Kinase website (http://www.kinase.com/web/current/), and the literature, and was organized manually, consolidated and checked for consistency.

B. DTO Visualization

Ontologies, as sets of concepts and their interrelations, are represented as graphs. Visualization of ontologies is important for the deep understanding of the underlying structure annotated in OWL ontology files. Existing visualizations for ontologies often focus on certain aspects of ontologies. The class hierarchy is visualized to facilitate with the graphical depictions of its main elements and relationship. The class concepts and subsumption (is-a) relationship are placed in a combination of force-directed, radial tree, hierarchical layouts, tree maps or their variants, for example, OWLViz [24], OntoTrack [30], GLOW [26], KC-Viz [27], OWL-VisMod [28], Jambalaya [29], OntoGraf [25], FlexViz [31], OLSVis [32], OWLPropViz [33], OntoSphere [34], Onto3DViz [35], and OntoSELF [36]. CropCircles [39] represents the class hierarchy as nested circles similar to treemaps. However, all these work are tailored to specific tasks. Different users' needs may call for different visualization models.

With the rapid increasing of the drug screening data at a large scale, there is an emerging need of examining, coordinating,

and integrating the data from diverse sources. Traditionally, drug design has been pursued under the philosophy, "one disease, one drug, one gene", where the primary objective is to find a drug compound that binds with high affinity to a target of interest. The philosophy has evolved into "magic shotguns" which suggests that many experimental efforts have been emphasized on multiple-target interactions. The importance of gaining selectivity is well understood. We appreciate existing approaches to construct selectivity screening panels. However, we are aware of the possibility that ligand specificity and selectivity in drug-target data is biased towards the hypothesis testing or depends on prior knowledge of the expertise.

The preference of ontology visualization relies on the need in examining the selectivity of diverse drug screening data with a goal of discovering the possible existence of the limited experimental exploration of 'dark matter' drug targets. We propose an ontology-based reference visualization model that facilitates with an intuitive view of the selectivity in a drug-target interaction network data. The overview procedure is shown in Figure 2.

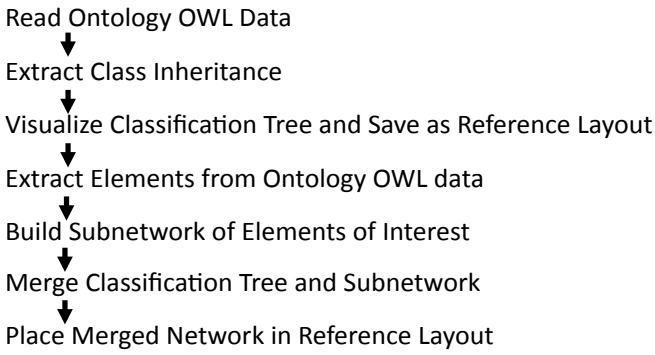

Figure 2. Overview of ontology-based reference visualization model.

In order to effectively analyze and illustrate the coverage of small molecules that bind to the human druggable genome, we employ DTO as the reference of druggable target knowledge base (version DTO_merged_v09_b3101), where a relatively complete collection of disease-linked human proteins and their belonging classification schema are enclosed. The classification of druggable proteins has been represented as in a tree-like hierarchical structure (also called subsumption tree), where leaf nodes represent as protein targets, non-leaf nodes as classes (also called categories), and edges adjacent to two nodes as the "is-a" subsumption relationship [14]. For example, ADRA1A is an alpha adrenoceptor (adrenergic receptor), which is a member of the GPCR superfamily that recognizes the endogenous ligand adrenaline and noradrenaline and which are amines and thus is also classifies as aminergic receptor [13]. In the subsumption "is-a" hierarchy, upper-level target class subsumes its lower-level target classes.

We have extracted and visualized the DTO subsumption tree by multiple runnings of the force-directed graph-drawing algorithm plugged in Cytoscape [11] (Figure 1A). The goal of the graph-drawing algorithm is to calculate layouts of nodes and edges of a graph in a planar surface. It uses repulsive forces between nodes and attractive forces between adjacent nodes. An objective function in the optimization algorithm is the calculation of the energy of the layout based on the forces of all pairwise nodes. There are diverse methods to calculate forces. In general, forces between the nodes are computed based on their graph theoretic distances, determined by the lengths of shortest paths between them. A best layout for a graph corresponds to finding a (often local) minimum of this objective function. As a consequence, in such a layout, low energies correspond to layouts in which adjacent nodes are near some pre-specified distance from each other, and in which non-adjacent nodes are well spaced [23]. We execute multiple runs till reaching to an aesthetically pleasant layout since each run may only find a local minimum under the default configuration in the Cytoscape, no matter how much iteration each run pursues. In the paper, we set up 1000 iterations for each run.

In the resulted visualization (see Figure 1), we illuminate different druggable protein families by colors. There are four colors, and each color stands for a superfamily (also called a IDG family, for example, G protein-coupled receptors (GPCRs), kinases, ion channels (ICs), and nuclear receptors (NRs)). Protein nodes are colored in light and protein class nodes in dark. We save the visualization as the reference layout model. In the reference graph, each node has the uniform size.

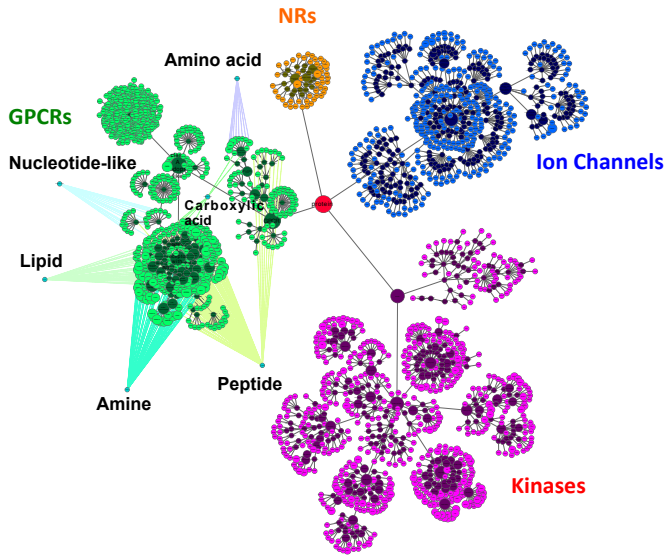

Figure 3. Visualization of DTO drug targets, its classification tree, and corresponding critical annotations. In the figure, six endogenous ligand types of GPCRs are shown (9 bind amino acids, 130 peptides, 29 lipids, 52 amines, 19 nucleotide-like ligands, and 6 GPCRs bind carboxylic acid ligands) (classification version by Aug. 26, 2015).

In addition to the "is-a" relationship, the DTO also included other relationships in order to characterize both chemical structural, and pattern-based or ligand-bound annotation of protein targets (Figure 3). For example, ADRA1A binds an amine endogenous ligand, or GRM5 is bound to an amino-

acid ligand. Since the classification of specific endogenous ligands for each G protein-coupled receptors (GPCRs) protein has been integrated with the source from IUPHAR and being mapped to the "Chemical entity" of the ChEBI database, it is meaningful to show the ligand classifications. We extracted the elements in DTO OWL file and built a subnetwork. Via the Cytoscape, we merged the subnetwork with the DTO classification tree and further visualize the unified network with the reference layout model. The approach can be used to visualize other properties of interest.

III. Visual Selectivity Analysis of Drug Screening Data

A. Drug Screening Data

Our interest in the integrative analysis has been inspired by the early successful replacement of "magic bullets" with "magic shortguns" in treating psychiatric diseases [9,10]. Among various databases, PDSP (NIMH Psychoactive Drug Screening Program) database is one of relatively complete and comprehensive publicly accessible databases, which authors have developed an automated, adaptive design approach to optimize ligands and extended the promiscuity among currently marked antipsychotic drugs for the Alzheimer's disease [1-2]. Our interest is how extensive the promiscuity has been profiled. Another widely used database is ChEMBL [6]. This database includes a large number of well-annotated literature-curated bioactivity data, including small molecule protein binding data [6]. We questioned the distribution of active protein targets in ChEMBL database, which are hit by all LINCS small molecules [8], with the suspicion that ligand specificity and selectivity in drug-target data is likely biased towards the hypothesis testing or depends on prior knowledge of the expertise, possibly showing the intensity of efforts from drug-target data source.

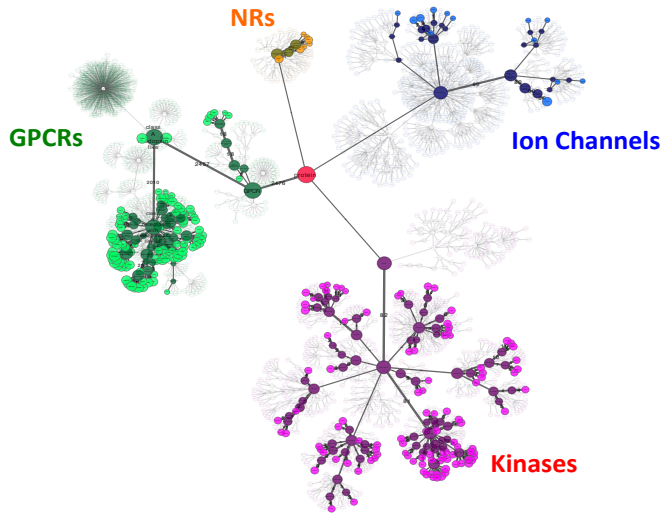

Figure 4. The coverage of active protein targets selected by PDSP compound library on DTO.

B. Ontology-based Visual Selectivity Analysis

For a drug-target data, we extracted the active drug-target interactions on the condition of a specified cutoff since our interest of selectivity examination lies in the on or off-targets profiled against compounds. The overview procedure is in Figure 5. We iteratively calculated active compounds $D(T_i)$ for a protein target T_i in the ontology, where active compounds are defined as those being involved in active drug-target interactions with T_i. Via the depth-first search algorithm in walking through all elements in the ontology classification tree, for a category class C_j in the ontology, we computed all belonging protein targets, $|T_i, \exists\, T_i \epsilon C_j|$, and aggregated active compounds, $|\bigcup_{T_i \epsilon C_j} D(T_i)|$, which hit or interact with at least one target in the class. We also employed the breadth-first search algorithm to identify the height H and $h(i)$, respectively for the classification tree and its each node i.

In the visual selectivity analysis, we used the reference ontology classification tree as the background and highlighted protein targets, interacting with active compounds, as well as their belonged classes and their location in the tree. Each node i is weighted by the scoring scheme $\log|T_i, \exists\, T_i \epsilon C_j| + \log\left|\bigcup_{T_i \epsilon C_j} D(T_i)\right| + \mathrm{h(i)}$. Each edge, representing a class subsumption (is-a) relationship , e.g., $C_i \rightarrow C_j$ (C_j is a special subclass of C_i), is weighted by active compounds interacting with targets belonging to C_j.

In Cytoscape, we updated node and edge attributes for the reference ontology model, labeled nodes by their IDs and labeled edges by their weight, specified node and edge opacity, and scaled nodes and edges by their weights.

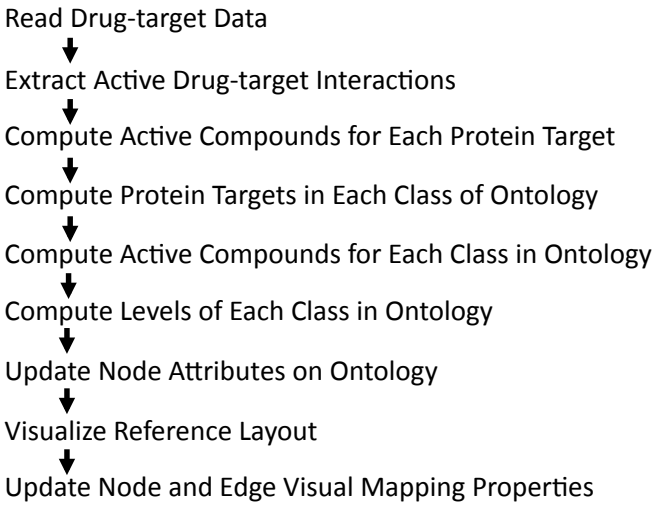

Figure 5. Overview of visualizing drug-target selectivity of screening data

C. Ontology-based Visual Selectivity Analysis on PDSP Data

We downloaded the NIMH Psychoactive Drug Screening Program (PDSP) data at [1-2], curated, performed a cross-reference to Uniprot IDs, and extracted the active drug-target combinations with a cutoff of $K_i < 10uM$, and then aggregated active compounds-hitting targets in each category. In order to

demonstrate the distribution of those compounds-hit targets in druggable proteome, we carried out the above-mentioned ontology-based visual selectivity analysis, using reference graph as the background with opacity = 0.5 and highlighting compound-interacting targets in PDSP and its belonging classes (see Figure 4). It shows the coverage of active protein targets selected by PDSP compound library on DTO.

D. Ontology-based Visual Selectivity Analysis on LINCS Compounds-interacting ChEMBL Database

We queried for the targets in ChEMBL database, which are hit by all LINCS small molecules. Through the further integrative analysis with DTO, we found 632 active drug targets inhibited by LINCS compounds (confidence value greater than or equal to 7 and aggregated ChEMBL value greater than or equal to 5). Though in total we have 1107 ChEMBL targets inhibited by LINCS Compounds, out of which we have observed a list of 552 PDSP compounds (688 PDSP samples) overlapping with LINCS compounds (by standardized SMILES). The "Overlapping" is represented as the case that the standardized smiles (after removing the salts and addends) are exactly the same. Standardization was done via PubChem standardization service. In the study, we used the same method to visualize the coverage of ChEMBL human druggable targets hit by the LINCS compound (see Figure 5).

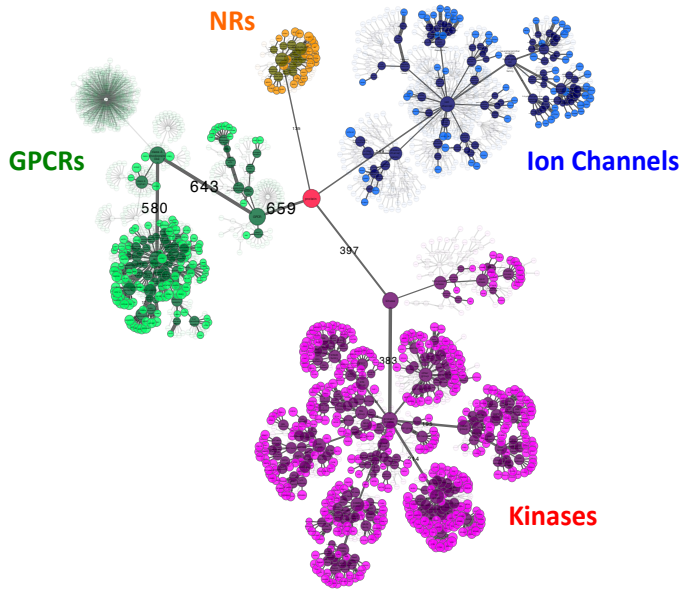

Figure 5. **DTO coverage of LINCS compound-hitting druggable ChEMBL targets.**

E. Discussion

The knowledge-driven ontology-based visual selectivity analysis illustrates areas with very limited experimental exploration, which may provide opportunities for new drug development efforts. For example, all 424 olfactory receptor proteins are not profiled in both screening data. It seems experimentally reasonable since there is the lack of comprehensive understanding of olfactory mechanism in relation to diverse chronic psychiatric diseases. However, it is noticed that olfaction plays an important role in protecting us from environmental dangers. Olfactory dysfunction is common in later life and early warning of neurodegenerative disease, which is receiving more attention as a sign in the early detection and differential diagnosis of neurological and mental diseases. The timely effective treatment of olfactory loss is facilitative to cure related complications of neurodegenerative disease.

Viewing from subfigure A and B in Figure 6, we observed that 1) there are a large number of compounds profiled to a limited number of GPCR proteins in PDSP data set, 2) in contradiction to the total druggable targets (gray bar in subgraph B), there still exist a large number of uncovered targets in both datasets. Details can be referred to Table 1.

Both Figure 6 and Table 1 illustrate the number of chemical ligands from PDSP and ChEMBL binding assays that are active against these four critical druggable protein target classes, demonstrating that current drug screening efforts are highly biased towards a small subset of likely druggable targets and that there still exist a large number of understudied druggable targets.

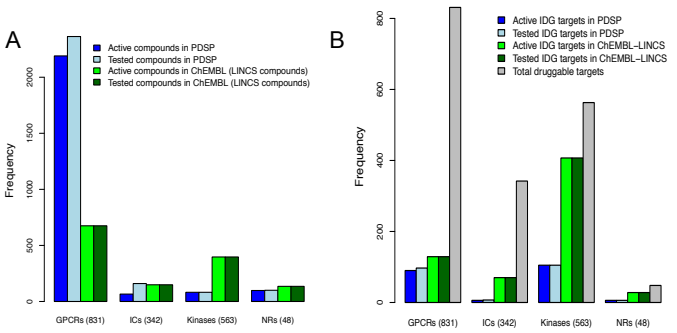

Figure 6. Experimental variance observed in PDSP data versus ChEMBL sub dataset that is inhibited by LINCS compounds and profiled for the Homo sapiens (with confidence value greater than or equal to 7 and standard type among Ki, Kd, IC50). **A.** The number of active compounds and tested compounds in two data sets, grouped by four important druggable protein target classes. **B.** The number of active targets and tested targets in two data sets, grouped by the same classes as A.

IDG families	Compounds				IDG Targets			
	PDSP		ChEMBL		PDSP		ChEMBL (driven by LINCS compounds)	
	Active compounds	Profiled compounds	Active LINCS compounds	Profiled LINCS compounds	Active targets	Profiled targets	Active targets	Profiled targets
GPCRs	2190	2363	675	735	90	97	129	134
Ion channels (ICs)	66	160	149	187	6	7	70	79
Kinases	82	82	397	446	105	105	407	407
Nuclear receptors (NRs)	98	100	135	165	6	6	28	28
Total	2419	2599	1204	1327	207	215	634	648

Table 1 Experimental variance in PDSP and ChEMBL drug screening data. For ChEMBL data, we extracted and aggregated a subset of ChEMBL data for homo sapiens, where LINCS compounds were tested with drug target binding confidence value greater than or equal to 7 and bioassay profile type were among Ki, Kd, and IC50. The active target or compound is defined if the ChEMBL value was greater than or equal to 5. For PDSP data, we added Human orthologs and extracted compound-target bioactivity profiles for homo sapiens and with meaningful identification. The active target or compound is defined if the PDSP Ki value was less than 10um. For both datasets, we grouped compounds, active compounds, targets, active targets by GPCRs, ICs, Kinases, NRs, where Uniprot IDs have been used. For each type, we examined the total number of targets and active targets tested in both data sets. Active targets are defined as being inhibited or bound with an aggregated value (greater than or equal to 5) by at least a tested compound in a specific data set. We also counted the total number of compounds and active compounds against each gene family in each data set.

IV. CONCLUSION

In a drug development process, appropriate drug-binding selectivity is critical for a success drug. However the selectivity in a data source may be limited to prior knowledge of the expertise or be biased towards the hypothesis testing. With the increasing and increasing of drug screening data, it is challenging to coordinate the efforts and execute data governance at a large scale. Visual selectivity analysis for examining target selection is in demand. In this paper, we introduced a novel method for the analysis of drug-target interactions based on DTO and visual analytics. We proposed a knowledge-driven approach and designed an ontology reference model to provide an intuitive view of the selectivity in a drug-target interaction network data when coordinating and fusing drug screening data. We examined the model on the NIMH Psychoactive Drug Screening Program (PDSP) data and the LINCS Compounds-interacting ChEMBL Database for the visual selectivity analysis. One of our other studies has validated the classification structure of DTO and further use DTO for similarity-based kernel-driven drug-target interaction prediction. The analysis in combination with biological regulatory process-based analysis helps indicate the possible 'dark matter' drug targets. The approach can be expanded to coordinate other experimental screening data and set a stage for the analysis of the mechanism of action of biological therapies.

V. ACKNOWLEDGEMENT

This work was supported by grants U54CA189205 (Illuminating the Druggable Genome Knowledge Management Center, IDG-KMC) and U54HL127624 (BD2K LINCS Data Coordination and Integration Center, DCIC). The IDG-KMC (http://targetcentral.ws/) is a component of the Illuminating the Druggable Genome (IDG) project (https://commonfund.nih.gov/idg) awarded by the NCI. The BD2K LINC DCIC is awarded by the National Heart, Lung, and Blood Institute through funds provided by the trans-NIH Library of Integrated Network-based Cellular Signatures (LINCS) Program (http://www.lincsproject.org/) and the trans-NIH Big Da-ta to Knowledge (BD2K) initiative (https://datascience.nih.gov/bd2k). Both IDG and LINCS are NIH Common Fund projects. Q. C., C. D., and C. C. have also been supported by the grant P031S150046, funded by the U.S. Department of Education.

REFERENCES

[1] Kroeze, W. K., Sassano, M. F., Huang, X. P., Lansu, K., McCorvy, J. D., Giguère, P. M., ... & Roth, B. L. (2015). PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nature structural & molecular biology*, *22*(5), 362-369.

[2] Besnard J, Ruda GF, Setola V, Abecassis K, Rodriguiz RM, Huang XP, Norval S, Sassano MF, Shin AI, Webster LA, Simeons FR, Stojanovski L, Prat A, Seidah NG, Constam DB, Bickerton GR, Read KD, Wetsel WC, Gilbert IH, Roth BL, Hopkins AL. (2012) Automated design of ligands to polypharmacological profiles. *Nature. 2012 Dec 13;492(7428):215-20. doi: 10.1038/nature11691.*

[3] Guilio Rastelli and Luca Pinzi. (2015) Computational polypharmacology comes of age. Frontiers in Pharmacology. Volume 6 | article 157. doi: 10.3389/fphar.2015.00157.

[4] Uhlén M et al, 2015. Tissue-based map of the human proteome. *Science* PubMed: 25613900 DOI: 10.1126/science.1260419

[5] Mestres, Jordi, et al. (2008) Data completeness—the Achilles heel of drug-target networks. *Nature biotechnology* 26.9: 983-984.

[6] A.P. Bento, A. Gaulton, A. Hersey, L.J. Bellis, J. Chambers, M. Davies, F.A. Kruger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos and J.P. Overington (2014) The ChEMBL bioactivity database: an update. Nucleic Acids Res., 42 1083-1090.

[7] Stephan C. Schürer and Steven M. Muskal (2012) Kinome-wide Activity Modeling from Diverse Public High-Quality Data Sets. J Chem Inf Model. 2013; 53(1): 27–38 DOI: 10.1021/ci300403k

[8] LINCS. http://lincs.hms.harvard.edu/db/

[9] Roth, B.L., Sheffler, D.J. & Kroeze, W.K. (2004) Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. Nat. Rev. Drug Discov. 3, 353–359.

[10] Wesley K. Kroeze and Bryan L. Roth. (2012) Polypharmacological Drugs: " Magic Shotguns " for Psychiatric Diseases. Peters, Jens-Uwe, ed. Polypharmacology in Drug Discovery. Somerset, NJ, USA: John Wiley & Sons.

[11] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., … Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research, 13(11), 2498–2504. http://doi.org/10.1101/gr.1239303

[12] Santos, R., Ursu, O., Gaulton, A., Bento, A.P., Donadi, R.S., Bologa, C., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T.I., & Overington, J.P. A Comprehensive Map of Molecular Drug Targets. Nature Rev. Drug Discov. doi:10.1038/nrd.2016.230 (2016)

[13] Przydzial, M. J., Bhhatarai, B., Koleti, A., Vempati, U., & Schürer, S. C. (2013). GPCR ontology: development and application of a G protein-coupled receptor pharmacology knowledge framework. Bioinformatics. 2013 Dec 15;29(24):3211-9. doi: 10.1093/bioinformatics/btt565. Epub 2013 Sep 29.

[14] Lin, Y., Mehta, S., McGinty, H. K., Turner, J. P., Vidovic, D., Forlin, M., ... & Mathias, S. L. (2017). Drug Target Ontology to Classify and Integrate Drug Discovery Data. bioRxiv, 117564.

[15] Cannon, D.C., Yang, J.J., Mathias, S.L., Ursu, O., Mani, S., Waller, A., Schürer, S.C., Jensen, L.J., Sklar, L.A., Bologa, C.G. and Oprea, T.I., 2017. TIN-X: target importance and novelty explorer. Bioinformatics, p.btx200.

[16] Ursu, O., Holmes, J., Knockel, J., Bologa, C.G., Yang, J.J., Mathias, S.L., Nelson, S.J. and Oprea, T.I., 2017. DrugCentral: online drug compendium. Nucleic acids research, 45(D1), pp.D932-D939.

[17] Loging, W.T. ed., 2016. Bioinformatics and Computational Biology in Drug Discovery and Development. Cambridge University Press.

[18] Acharya, A. and Garg, L.C., 2016. Drug Target Identification and Prioritization for Treatment of Ovine Foot Rot: An In Silico Approach. International journal of genomics, 2016

[19] Russ, A. P., & Lampel, S. (2005). The druggable genome: an update. Drug discovery today, 10(23), 1607-1610.

[20] Kolaja, K. (2012). Drug discovery: computer model predicts side effects. Nature, 486(7403), 326-328.

[21] Lounkine, E., Keiser, M.J., Whitebread, S., Mikhailov, D., Hamon, J., Jenkins, J., Lavan, P., Weber, E., Doak, A.K., Côté, S. and Shoichet, B.K., 2012. Large scale prediction and testing of drug activity on side-effect targets. Nature, 486(7403), p.361.

[22] Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. British journal of pharmacology, 162(6), 1239-1249.

[23] Kobourov, S. G. (2004). Force-directed drawing algorithms.

[24] M. Horridge. OWLViz. http://protegewiki.stanford.edu/wiki/OWLViz, 2010.

[25] Liebig, T., & Noppens, O. (2005). OntoTrack: A semantic approach for ontology authoring. Web Semantics: Science, Services and Agents on the World Wide Web, 3(2), 116-131.

[26] Hop, W., de Ridder, S., Frasincar, F., & Hogenboom, F. (2012, March). Using Hierarchical Edge Bundles to visualize complex ontologies in GLOW. In Proceedings of the 27th Annual ACM Symposium on Applied Computing (pp. 304-311). ACM.

[27] Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. Scientific american, 284(5), 28-37.

[28] García-Peñalvo, F. J., Colomo-Palacios, R., García, J., & Therón, R. (2012). Towards an ontology modeling tool. A validation in software engineering scenarios. Expert Systems with Applications, 39(13), 11468-11478.

[29] Storey, M. A., Noy, N. F., Musen, M., Best, C., Fergerson, R., & Ernst, N. (2002, January). Jambalaya: an interactive environment for exploring ontologies. In Proceedings of the 7th international conference on Intelligent user interfaces (pp. 239-239). ACM.

[30] S. Falconer. OntoGraf. http://protegewiki.stanford.edu/wiki/OntoGraf, 2010.

[31] Falconer, S. M., Callendar, C., & Storey, M. A. (2010, October). A visualization service for the semantic web. In International Conference on Knowledge Engineering and Knowledge Management (pp. 554-564). Springer, Berlin, Heidelberg.

[32] Vercruysse, S., Venkatesan, A., & Kuiper, M. (2012). OLSVis: an animated, interactive visual browser for bio-ontologies. BMC bioinformatics, 13(1), 116.

[33] L. Wachsmann. OWLPropViz. http://protegewiki.stanford.edu/wiki/OWLPropViz, 2008.

[34] Bosca, A., Bonino, D., & Pellegrino, P. (2005, December). OntoSphere: more than a 3D ontology visualization tool. In Swap.

[35] Guo, S. S., & Chan, C. W. (2010, May). A tool for ontology visualizaiton in 3D graphics: Onto3DViz. In Electrical and Computer Engineering (CCECE), 2010 23rd Canadian Conference on (pp. 1-4).

[36] M. Sintek. OntoViz. http://protegewiki.stanford.edu/wiki/OntoViz, 2007.

[37] Eklund, P., Roberts, N., & Green, S. (2002). Ontorama: Browsing rdf ontologies using a hyperbolic-style browser. In Cyber Worlds, 2002. Proceedings. First International Symposium on (pp. 405-411). IEEE