Eurographics Workshop on Visual Computing for Biology and Medicine (2024)
L. Garrison and D. Jönsson (Short Paper Chairs)

Short Paper

VISPER - Visualization System for Interactions between Proteins and Drugs for Exploratory Research

D. Dehncke[1,2,†], V. Fiebach[1,2,†], L. Kinzel[3], K. Baumann[3], T. Kacprowski[1,2]

[1]Division Data Science in Biomedicine, Peter L. Reichertz Institute for Medical Informatics, Technische Universität Braunschweig and Hannover Medical School
[2]Braunschweig Integrated Centre of Systems Biology (BRICS), Technische Universität Braunschweig
[3]Institute of Medicinal and Pharmaceutical Chemistry, Technische Universität Braunschweig
[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint first authors.

Abstract

VISPER is a web-based application that enables users to interactively explore and analyze drug-protein associations. Its uniqueness lies in the dataset for which it has been specifically designed. Until now, most biomarkers for cancer vulnerabilities have primarily relied on genomic and transcriptomic measurements. A recently published study created a comprehensive pan-cancer proteomic map of human cancer cell lines, involving the application of 625 drugs to these cell lines. From these data, proteomic responses to the drug treatment across different cell lines can be derived, providing an extensive resource for a better understanding of drug mechanisms. To facilitate the analysis of this extensive dataset, we developed VISPER, a visualization tool specifically tailored to explore the ProCan dataset, enabling easy exploration of the relationships between proteins, drugs, and cell lines through a network graph representation. The graphical representation is complemented by a wide range of filter options, different representations, and integration of existing online databases for improved biological classification. Furthermore, the web application provides a clear overview of the similarity of drugs based on their protein associations. VISPER thus represents a promising addition to established systems biology software tools.

Availability and implementation: *VISPER is available open-source on GitHub (https://github.com/scibiome/VISPER) or as a Docker image (https://hub.docker.com/r/thegoldenphoenix/VISPER).*

CCS Concepts

• ***Human-centered computing*** → *Visual Analytics;* • ***Applied computing*** → *Life and medical sciences;*

1. Introduction

In the last decades advancements in drug research have been largely driven by high-throughput technologies, enabling the generation of extensive biological data at relatively low costs. Publicly available datasets in the domain of cancer research include the NCI-60 database [Sho06] or the Genomics of Drug Sensitivity in Cancer (GDSC) project [IKV*16]. These databases have shown that pharmacogenomic profiling of cancer cell lines derived from clinical tumor samples can significantly aid in the development of new cancer therapies [YSG*13].

Most oncological biomarkers are based on genomic or transcriptomic profiles [GHJV*19]. However, transcript levels alone cannot adequately predict protein levels or explain genotype-phenotype relationships [LBA16]. The complexity of cancer and the poor performance of genomics and transcriptomics in predicting the abundance and activity of proteins suggests that these markers alone are not sufficient to adequately guide clinical care for many patients.

Until recently, proteomics quantification in cancer cell lines was limited by rather low numbers and narrow focus on specific cancer types [GPC*22]. Consequently, the proteome's contribution to cancer vulnerabilities in different tissue and genetic contexts is not well studied. By using a data-independent acquisition (DIA) mass spectrometry (MS) approach combined with novel data processing methods, it is now possible to achieve reproducible proteomics at scale [PHS*20]. Creation and dissemination of extensive proteomic datasets open up opportunities for developing new computational methods, to explore how molecular changes affect cancer vulnerabilities.

Such large datasets have proven useful in many fields, even beyond oncology. This is especially the case if there exist approaches to (semi-)automatically mine them. The process of mining extensive datasets greatly benefits from user-friendly interfaces, standardized data structures, and computational biology tools to support researchers. Tools like *GDSCTools* [CCI*18] have proven helpful by incorporating statistical tools, predictive methods, and the capability to implement additional analytical methods, providing common data structures.

DOI: 10.2312/vcbm.20241190

In a recent study, Gonçalves et al. [GPC*22] created a pan-cancer proteome map, aiming to aid in discovering cancer biomarkers and targets for new treatments. The dataset, which we will call ProCan henceforth, facilitates the correlation of drug response data with proteomic profiles. The study found that the predictive power of proteomics and transcriptomics is comparable, with evidence that proteomics can provide additional crucial information not covered by transcriptomics.

To make the ProCan dataset accessible to a wider audience, the "Bio+MedVis Challenge @ IEEE VIS 2023" was organized [noaa]. This initiative aims to develop tools allowing users to explore the dataset, uncover hidden patterns, and perform analyses that would be difficult or time-consuming with the raw data. Based on the requirements of this challenge, we have developed the visualization tool "VIsualization System for Interactions between Proteins and Drugs for Exploratory Research" (VISPER). With VISPER, users can search for individual entities or groups of entities (e.g., drugs, proteins, or cell lines) and display relationships with other entities. In addition to finding connections and filtering results, VISPER offers the option to upload custom data and export results. Our application also identifies drugs with similar effects, which can aid in searching for candidates for drug repurposing.

2. Related Work

Analyzing biological data presents significant challenges, particularly in identifying discerning patterns and relationships within complex datasets. Network visualization plays a crucial role in this process, but the choice of an appropriate visualization method depends heavily on the type of network being studied [KKPEP20].

General-purpose tools like Pajek [MB16] are versatile for broad analytical tasks and can be adapted for biological network analysis. Specialized tools such as Cytoscape offer focused capabilities in network visualization and annotation, supported by extensive plugin support that enhances its functionality [SMO*03]. Cytoscape.js [FLH*16], a JavaScript library inspired by the desktop application, provides similar functionalities.

Most omics databases provide basic visualization and analysis methods, such as BioGRID [OSB*19], which focuses primarily on protein-protein interactions. In the realm of drug development and repurposing, Drugst.One [MHA*24] provides a broad platform for modeling and analyzing complex drug-protein-disease networks.

In the context of the "Bio+MedVis Challenge @ IEEE VIS 2023", CytoCave [CKA*] facilitates the exploration of protein associations and drug similarities in the ProCan dataset. It emphasizes global visualization through the use of dimensional reduction and community detection techniques.

Existing tools do not adequately address the complexity of the ProCan dataset, or do not integrate essential external databases required for contextualizing the findings. This necessitates the development of a custom-made tool capable of thoroughly exploring the dataset without compromising the depth of analysis.

3. Material & Methods

The dataset provided by Gonçalves et al. [GPC*22] comprises 949 cancer cell lines representing 28 tissue types, accompanied by data on 625 drugs and over 40 cancer types. Each cell line's proteome was quantified, yielding an median of 5,237 proteins per cell line. This extensive dataset significantly advances the molecular characterization of specific cancer cell line models [GPC*22]. The authors provide two distinct datasets, GDSC1 and GDSC2, derived from different measurement methods. While these datasets overlap in content, they can be explored separately within the provided application to facilitate a more nuanced analysis.

Of the tasks presented in the "Bio+MedVis Challenge @ IEEE VIS 2023" we focused specifically on the third task, which can be summarized into the following key objectives:

- Create an interactive tool to visualize drug response data across cell lines in relation to their proteomic data.
- Allow users to select a protein and analyse drugs that target them
- Allow users to select a specific drug and visualize proteins that are strongly associated with the drug response

Our tool enables exploration across multiple entities and facilitates refined visualization and analysis. Selections can be filtered based on multiple criteria, allowing examination not only of drug-protein associations but also at the cell line level. Our integrated databases contextualize acquired knowledge within a broader framework. Additionally, we have introduced further functionalities to enhance analysis, often drawing inspiration from tools such as Drugst.One [MHA*24] and CytoCave [CKA*]. In contrast to CytoCave, our approach prioritizes exploratory analysis at a more localized interaction level.

3.1. Design and Development of *VISPER*

We have developed a platform-independent web application accessible via the provided Docker container [Mer]. Our backend utilizes Neo4j [noad], a graph database management system known for handling large network data and offering functions like shortest path calculations between nodes. VISPER uses FastAPI [noac] as server framework, connecting the Neo4j database to the frontend.

Given the complexity of our dataset and the diversity of nodes and edges, we employed Cytoscape.js [FLH*16], a JavaScript library tailored for graph visualization and analysis. Cytoscape.js offers a variety of extensions and provides a range of layout options. We created interactive plots that allow comprehensive exploration of the dataset, enabling users to select specific entities for focused analysis, such as examining drugs targeting particular pathways or exploring cell lines by tissue type. To achieve this, we utilized D3 [noab] and Plotly.js [noae]. For investigating proteins beyond those directly measured in ProCan, we integrated data from the BioGRID database (version 4.4.277) [OSB*19].

In Figure 1, the VISPER user interface is presented. The right hand side shows the results of the user query as graph. On the left, detailed information about the selected node is provided, along with dedicated links for further exploration, various filter options, and a range of analysis methods. These methods include:

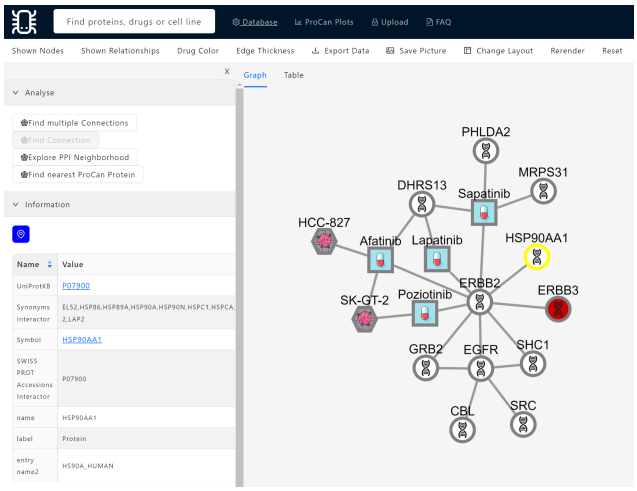

Figure 1: *The VISPER user interface features a main panel showing associations between proteins (circles), drugs (squares), and cancer cell lines (hexagons). The selected protein is highlighted in yellow, while proteins with a red background are not in the ProCan dataset but have been included based on connections from BioGRID. On the left side, information about the selected node is provided, along with filtering options and additional analysis tools. Users can configure the layout options at the top.*

- Find Connections: Select one or more nodes to find connections to other nodes based on specified criteria, such as identifying drugs associated with a source protein or exploring protein-protein interactions. The 'Find Connections' functionality fulfils the challenge requirements by identifying all proteins associated with a specific drug, and vice versa.
- Shortest Path: Finds the shortest path between two selected nodes. Additionally, the user can specify which kind of connections are considered to find the path.
- Explore PPI Neighborhood: Identifies proteins based on their centrality in the BioGRID PPI network relative to the selected proteins. Allows to find proteins that are interacting with the given set of proteins.
- Find Nearest ProCan Protein: Offers a function to select a BioGRID protein and identify the nearest ProCan protein. This approach helps finding interacting proteins and potential drug associations.

To facilitate further analysis, we provide a wide range of outbound links to domain-specific websites and databases. For proteins, we integrate links to UniProt [The15]. Selecting specific proteins also allows users to explore their interactions on external platforms such as Drugst.One [MHA*24]. For drugs, we offer links to DrugBank [WFG*18].

Additionally, we incorporated data from DrugBank [WFG*18] and ChEMBL [ZFH*24]. DrugBank provides comprehensive molecular information about drugs, including their mechanisms of action, drug interactions, and targets. ChEMBL is a curated database of bioactive molecules with drug-like properties, integrating chemical, bioactive, and genomic data.

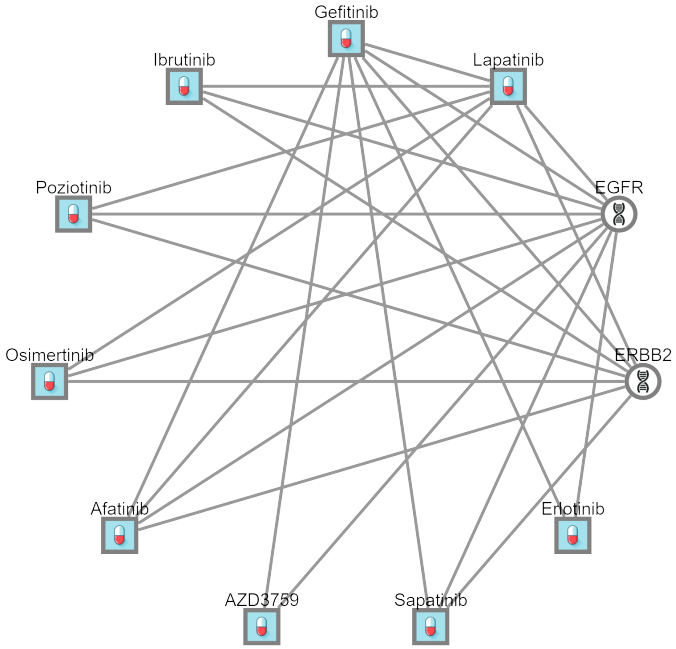

Figure 2: *Visualization of the associations between various TKIs and the proteins EGFR and ERBB2. Starting with Gefitinib and Lapatinib, similar drugs were identified using the drug similarity calculated by VISPER. Edges between the quadratic drug nodes can be examined by the user for information about the exact similarity score. Likewise, if an edge between a circular protein node and a drug is selected, VISPER shows parameters and statistics of the regression model on which the association is based. Notably, the graph is not fully connected because a filter was applied to only show drug-protein associations within an 1% FDR threshold.*

4. Use Cases

Use case 1: Finding information about specific drug-protein associations and pathways of interest

In their publication, Gonçalves et al. [GPC*22] showcased significant associations between protein intensity and drug response that they found for EGFR related proteins and the tyrosine kinase inhibitors (TKIs) Gefitinib and Lapatinib. In order to demonstrate how VISPER can be used to explore connections of interest, we investigated the associations of several other TKIs with EGFR and ERBB2 to show if the findings align with the current literature. We started by adding Gefitinib, Lapatinib, EGFR and ERBB2 to the graph, utilizing the search function. Using VISPER's 'Find multiple connections' analysis function, we connected both Gefitinib and Lapatinib to their five most similar drugs. Like in previous work from Chukhman et al. [CKA*], we calculated drug similarity as the cosine similarity over the beta values of the shared associations of two drugs. A higher similarity thereby indicates, that the response to both drugs is similarly affected by the cellular proteome. The resulting graph contained the two proteins, as well as nine drugs, all of which were listed by Yang et al. [YSG*13] to have EGFR among their targets, except for Ibrutinib, which is commonly classified as an inhibitor of Bruton's tyrosine kinase, al-

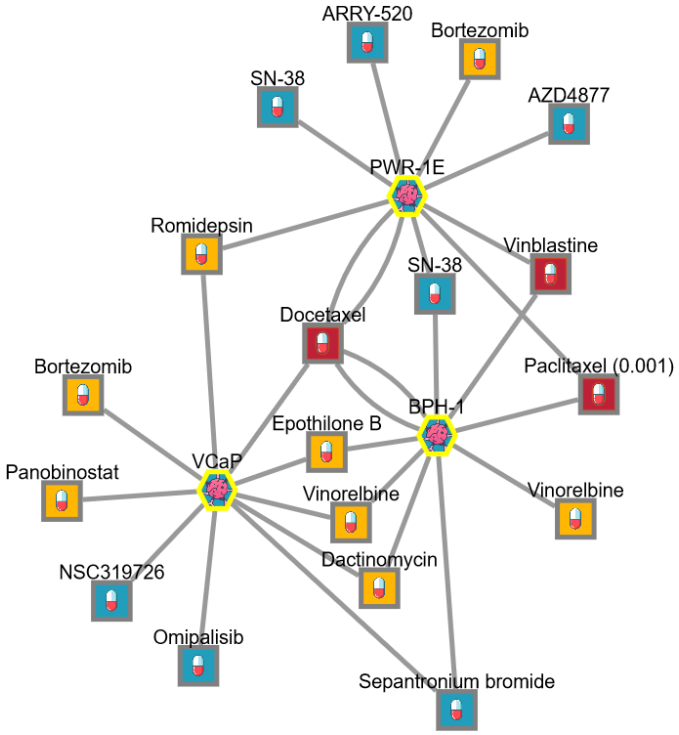

Figure 3: *A graph created with VISPER in order to find drugs that inhibit the growth of VCaP, BPH-1 and PWR-1E, which are used in prostate cancer research. Each cell line was connected to its ten inhibitors with the lowest ln(IC_{50}) values. Notably, Docetaxel shares two edges with BPH-1 and PWR-1E each because the results include associations from GDSC1 and GDSC2. Docetaxel also shares an edge with VCaP and can therefore be easily identified to have a high efficacy on all three cell lines. The color coding used in this graph shows all compounds in yellow, that were used in phase III clinical trials related to cancer or are already approved according to the ChEMBL database. Similarly, compounds in red were tested in phase III against prostate cancer or are already approved for prostate cancer.*

though it also inhibits EGFR at low nanomolar concentrations due to structural homology [HPZ*18]. All drug nodes were connected to ERBB2 and EGFR using the 'Find connection' function, to find the respective drug response – protein measurement association from GDSC2 for each possible drug-protein combination. As recommended by Gonçalves et al., the resulting connections were filtered to have a false discovery rate (FDR) below 1%. The resulting graph is shown in Fig. 2. Clicking on the connections in the graph, it can easily be seen by the user, that all TKIs in the graph indeed show similar negative associations with EGFR. Further, all significant associations of ERBB2 were with drugs that are either known to be strong inhibitors of ERBB2, to specifically inhibit the growth of cancer cells overexpressing ERBB2 [CKS*16] or to be at least weakly active against ERBB2, as in the case of Osimertinib [ETWA21]. Vice versa, all three drugs that don't share an association below the FDR threshold with ERBB2, have EGFR listed as their main target [YSG*13]. Notably, in this graph all associations with EGFR were weaker than those with ERBB2, which is in line with previous studies that found EGFR levels to be a poor predictor of TKI sensitivity [KNH*03] compared to ERBB2 expression [CKvG*21].

Use case 2: Navigating the GDSC database in search for multiple drug-cell line connections under specific criteria

The default web interface of the GDSC database [YSG*13] currently allows for a comprehensive overview of information associated with single compounds and individual cell lines. VISPER complements this functionality by providing a visual and intuitive interface to find connections with specified properties between multiple different entities. Consider the task of finding a small set of approved drugs, that inhibit the growth of specific cell lines with a minimum potency. This could be of interest for a research group that is planning the experimental setup of a small phenotypic screening and wants to select compounds to use as a positive control to assess the validity of their results. In the first step, the cell lines used in the screening could be added to the VISPER graph, for instance by using the list search function. In the second step, the most potent drugs could be connected to each cell line based on individual ln(IC_{50}) values using the 'Find multiple connections' button. After that, compounds that are highly active on multiple cell lines can be visually identified. Lastly, the color coding could be used to highlight compounds that are already approved and therefore likely to be commercially more easily accessible. A detailed example with the resulting visualization is given in Fig. 3. Further information on the cell lines or the identified drugs as well as the option to filter for ln(IC_{50}) value ranges of interest are provided by VISPER during each step through a comprehensive toolbar.

5. Conclusions and Future Work

We developed VISPER, an interactive visualization tool specifically designed for the detailed analysis of drug-protein interactions within the ProCan dataset. Addressing the third task of the "Bio+MedVis Challenge @ IEEE VIS 2023", VISPER includes advanced functionalities that enable researchers to conduct comprehensive analyses and contextualize their findings using external databases. Through empirical evaluations, we demonstrate VISPER's efficacy in supporting complex investigations into the ProCan dataset. Our future research aims to incorporate capabilities for assessing the chemical similarities between novel drug candidates and established compounds, thereby enhancing our ability to predict their potential protein interactions.

References

[CCI*18] COKELAER T., CHEN E., IORIO F., MENDEN M. P., LIGHTFOOT H., SAEZ-RODRIGUEZ J., GARNETT M. J.: GDSCTools for mining pharmacogenomic interactions in cancer. *Bioinformatics 34*, 7 (Apr. 2018), 1226–1228. doi:10.1093/bioinformatics/btx744. 1

[CKA*] CHUKHMAN M., KEFI A., AJILORE O. A., LEOW A., FORBES A. G.: CytoCave: An Interactive Visualization Tool for Exploring Protein and Drug Interaction Networks. URL: https://virtual.ieeevis.org/year/2023/paper_a-biomedchallenge-7741.html. 2, 3

[CKS*16] CHEN J., KINOSHITA T., SUKBUNTHERNG J., CHANG B. Y., ELIAS L.: Ibrutinib Inhibits ERBB Receptor Tyrosine Kinases and HER2-Amplified Breast Cancer Cell Growth. *Molecular Cancer Therapeutics 15*, 12 (Dec. 2016), 2835–2844. doi:10.1158/1535-7163.MCT-15-0923. 4

[CKvG*21] CONLON N. T., KOOIJMAN J. J., VAN GERWEN S. J. C., MULDER W. R., ZAMAN G. J. R., DIALA I., ELI L. D., LALANI A. S., CROWN J., COLLINS D. M.: Comparative analysis of drug response and gene profiling of HER2-targeted tyrosine kinase inhibitors. *British Journal of Cancer 124*, 7 (Mar. 2021), 1249–1259. doi:10.1038/s41416-020-01257-x. 4

[ETWA21] EWER M. S., TEKUMALLA S. H., WALDING A., ATUAH K. N.: Cardiac Safety of Osimertinib: A Review of Data. *Journal of Clinical Oncology 39*, 4 (Feb. 2021), 328–337. doi:10.1200/JCO.20.01171. 4

[FLH*16] FRANZ M., LOPES C. T., HUCK G., DONG Y., SUMER O., BADER G. D.: Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics 32*, 2 (Jan. 2016), 309–311. doi:10.1093/bioinformatics/btv557. 2

[GHJV*19] GHANDI M., HUANG F. W., JANÉ-VALBUENA J., KRYUKOV G. V., LO C. C., MCDONALD III E. R., BARRETINA J., GELFAND E. T., BIELSKI C. M., LI H., ET AL.: Next-generation characterization of the cancer cell line encyclopedia. *Nature 569*, 7757 (2019), 503–508. 1

[GPC*22] GONÇALVES E., POULOS R. C., CAI Z., BARTHORPE S., MANDA S. S., LUCAS N., BECK A., BUCIO-NOBLE D., DAUSMANN M., HALL C., HECKER M., KOH J., LIGHTFOOT H., MAHBOOB S., MALI I., MORRIS J., RICHARDSON L., SENEVIRATNE A. J., SHEPHERD R., SYKES E., THOMAS F., VALENTINI S., WILLIAMS S. G., WU Y., XAVIER D., MACKENZIE K. L., HAINS P. G., TULLY B., ROBINSON P. J., ZHONG Q., GARNETT M. J., REDDEL R. R.: Pan-cancer proteomic map of 949 human cell lines. *Cancer cell 40*, 8 (2022), 835–849.e8. doi:10.1016/j.ccell.2022.06.010. 1, 2, 3

[HPZ*18] HE L., PEI H., ZHANG C., SHAO M., LI D., TANG M., WANG T., CHEN X., XIANG M., CHEN L.: Design, synthesis and biological evaluation of 7*H*-pyrrolo[2,3-*d*]pyrimidin-4-amine derivatives as selective Btk inhibitors with improved pharmacokinetic properties for the treatment of rheumatoid arthritis. *European Journal of Medicinal Chemistry 145* (Feb. 2018), 96–112. doi:10.1016/j.ejmech.2017.12.079. 4

[IKV*16] IORIO F., KNIJNENBURG T. A., VIS D. J., BIGNELL G. R., MENDEN M. P., SCHUBERT M., ABEN N., GONÇALVES E., BARTHORPE S., LIGHTFOOT H., ET AL.: A landscape of pharmacogenomic interactions in cancer. *Cell 166*, 3 (2016), 740–754. 1

[KKPEP20] KOUTROULI M., KARATZAS E., PAEZ-ESPINO D., PAVLOPOULOS G. A.: A Guide to Conquer the Biological Network Era Using Graph Theory. *Frontiers in Bioengineering and Biotechnology 8* (Jan. 2020). Publisher: Frontiers. doi:10.3389/fbioe.2020.00034. 2

[KNH*03] KRIS M. G., NATALE R. B., HERBST R. S., LYNCH THOMAS J. J., PRAGER D., BELANI C. P., SCHILLER J. H., KELLY K., SPIRIDONIDIS H., SANDLER A., ALBAIN K. S., CELLA D., WOLF M. K., AVERBUCH S. D., OCHS J. J., KAY A. C.: Efficacy of Gefitinib, an Inhibitor of the Epidermal Growth Factor Receptor Tyrosine Kinase, in Symptomatic Patients With Non–Small Cell Lung CancerA Randomized Trial. *JAMA 290*, 16 (Oct. 2003), 2149–2158. doi:10.1001/jama.290.16.2149. 4

[LBA16] LIU Y., BEYER A., AEBERSOLD R.: On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell 165*, 3 (Apr. 2016), 535–550. doi:10.1016/j.cell.2016.03.014. 1

[MB16] MRVAR A., BATAGELJ V.: Analysis and visualization of large networks with program package Pajek. *Complex Adaptive Systems Modeling 4*, 1 (Dec. 2016), 6. doi:10.1186/s40294-016-0017-8. 2

[Mer] MERKEL D.: Docker: Lightweight Linux Containers for Consistent Development and Deployment. 2

[MHA*24] MAIER A., HARTUNG M., ABOVSKY M., ADAMOWICZ K., BADER G. D., BAIER S., BLUMENTHAL D. B., CHEN J., ELKJAER M. L., GARCIA-HERNANDEZ C., ET AL.: Drugst. one—a plug-and-play solution for online systems medicine and network-based drug repurposing. *Nucleic Acids Research* (2024), gkae388. 2, 3

[noaa] Bio+MedVis Challenge @ IEEE VIS 2023 – Program. URL: http://biovis.net/2023/program_ieee/. 2

[noab] D3. URL: https://d3js.org/. 2

[noac] FastAPI: Concurrency and async / await. URL: https://fastapi.tiangolo.com/async/. 2

[noad] Neo4j: Cypher Manual - Introduction. URL: https://neo4j.com/docs/cypher-manual/current/introduction/. 2

[noae] plotly.js. URL: https://www.npmjs.com/package/plotly.js. 2

[OSB*19] OUGHTRED R., STARK C., BREITKREUTZ B.-J., RUST J., BOUCHER L., CHANG C., KOLAS N., O'DONNELL L., LEUNG G., MCADAM R., ZHANG F., DOLMA S., WILLEMS A., COULOMBE-HUNTINGTON J., CHATR-ARYAMONTRI A., DOLINSKI K., TYERS M.: The BioGRID interaction database: 2019 update. *Nucleic acids research 47*, D1 (2019), D529–D541. doi:10.1093/nar/gky1079. 2

[PHS*20] POULOS R. C., HAINS P. G., SHAH R., LUCAS N., XAVIER D., MANDA S. S., ANEES A., KOH J. M., MAHBOOB S., WITTMAN M., ET AL.: Strategies to enable large-scale proteomics for reproducible research. *Nature communications 11*, 1 (2020), 3793. 1

[Sho06] SHOEMAKER R. H.: The NCI60 human tumour cell line anticancer drug screen. *Nature Reviews Cancer 6*, 10 (Oct. 2006), 813–823. Publisher: Nature Publishing Group. doi:10.1038/nrc1951. 1

[SMO*03] SHANNON P., MARKIEL A., OZIER O., BALIGA N. S., WANG J. T., RAMAGE D., AMIN N., SCHWIKOWSKI B., IDEKER T.: Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research 13*, 11 (Nov. 2003), 2498–2504. doi:10.1101/gr.1239303. 2

[The15] THE UNIPROT CONSORTIUM: UniProt: a hub for protein information. *Nucleic Acids Research 43*, D1 (Jan. 2015), D204–D212. doi:10.1093/nar/gku989. 3

[WFG*18] WISHART D. S., FEUNANG Y. D., GUO A. C., LO E. J., MARCU A., GRANT J. R., SAJED T., JOHNSON D., LI C., SAYEEDA Z., ASSEMPOUR N., IYNKKARAN I., LIU Y., MACIEJEWSKI A., GALE N., WILSON A., CHIN L., CUMMINGS R., LE D., PON A., KNOX C., WILSON M.: DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Research 46*, D1 (2018), D1074–D1082. doi:10.1093/nar/gkx1037. 3

[YSG*13] YANG W., SOARES J., GRENINGER P., EDELMAN E. J., LIGHTFOOT H., FORBES S., BINDAL N., BEARE D., SMITH J. A., THOMPSON I. R., RAMASWAMY S., FUTREAL P. A., HABER D. A., STRATTON M. R., BENES C., MCDERMOTT U., GARNETT M. J.: Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Research 41*, D1 (Jan. 2013), D955–D961. doi:10.1093/nar/gks1111. 1, 3, 4

[ZFH*24] ZDRAZIL B., FELIX E., HUNTER F., MANNERS E. J., BLACKSHAW J., CORBETT S., DE VEIJ M., IOANNIDIS H., LOPEZ D. M., MOSQUERA J. F., ET AL.: The chembl database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic acids research 52*, D1 (2024), D1180–D1192. 3