Cancer driver drug interaction explorer

Michael Hartung [1], **Elisa Anastasi**[2], **Zeinab M. Mamdouh** [3,4], **Cristian Nogales** [3],
Harald H.H.W. Schmidt[3], **Jan Baumbach**[1,5], **Olga Zolotareva** [1,6,†] **and Markus List** [6,*,†]

[1]Institute for Computational Systems Biology, University of Hamburg, 22607 Hamburg, Germany, [2]School of Computing, Newcastle University, 2308 Newcastle upon Tyne, UK, [3]Department of Pharmacology and Personalised Medicine, Maastricht University, 6229 Maastricht, Netherlands, [4]Department of Pharmacology and Toxicology, Faculty of Pharmacy, Zagazig University, 44519 Zagazig, Egypt, [5]Computational Biomedicine Lab, Department of Mathematics and Computer Science, University of Southern Denmark, 5230 Odense, Denmark and [6]Chair of Experimental Bioinformatics, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany

ABSTRACT

Cancer is a heterogeneous disease characterized by unregulated cell growth and promoted by mutations in cancer driver genes some of which encode suitable drug targets. Since the distinct set of cancer driver genes can vary between and within cancer types, evidence-based selection of drugs is crucial for targeted therapy following the precision medicine paradigm. However, many putative cancer driver genes can not be targeted directly, suggesting an indirect approach that considers alternative functionally related targets in the gene interaction network. Once potential drug targets have been identified, it is essential to consider all available drugs. Since tools that offer support for systematic discovery of drug repurposing candidates in oncology are lacking, we developed CADDIE, a web application integrating six human gene-gene and four drug-gene interaction databases, information regarding cancer driver genes, cancer-type specific mutation frequencies, gene expression information, genetically related diseases, and anticancer drugs. CADDIE offers access to various network algorithms for identifying drug targets and drug repurposing candidates. It guides users from the selection of seed genes to the identification of therapeutic targets or drug candidates, making network medicine algorithms accessible for clinical research. CADDIE is available at https://exbio.wzw.tum.de/caddie/ and programmatically via a python package at https://pypi.org/project/caddiepy/.

GRAPHICAL ABSTRACT
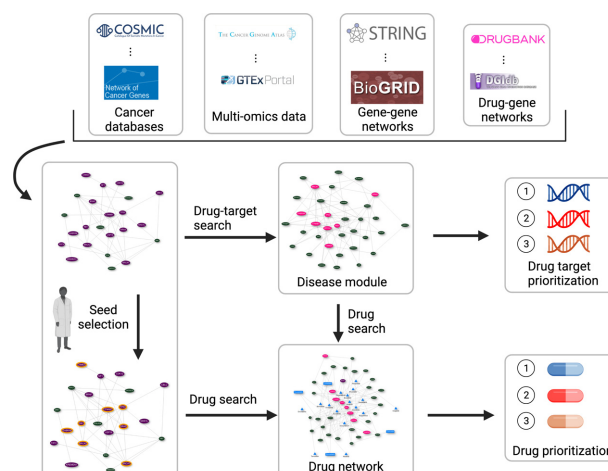
INTRODUCTION

Cancer is a set of diseases caused by alterations of the genome that lead to unregulated cell growth. Cancer therapy development is challenged by the heterogeneity of the disease, where similar phenotypes show distinct somatic mutations (1). Conversely, different cancer entities often have shared molecular mechanisms that can be exploited in network pharmacology (2). Since anti-cancer drugs cause extensive side effects, it is important to target molecular lesions and affected pathways specific to a tumor (3,4) and to adapt treatment in response to drug resistance (1). To identify suitable drugs, classical drug development is unsuitable as it is time consuming, expensive and often terminates in failure (5). A fast and cost-effective alternative to drug development is drug repurposing, the search of potentially effective drugs among the drugs approved for other indications (3).

Drug repurposing candidates can be identified computationally, experimentally or synergistically, where computational approaches lead to a hypothesis which can be validated experimentally (6). Network-based algorithms have proven successful for drug repurposing and to detect possible off-target effects (7). Computational drug repurposing methods based on protein-protein interaction (PPI) networks have been applied to find possible drugs against SARS-CoV-2 (8,9). In cancer, network-based repurposing allows to predict and prioritize drugs particularly for a group of cancer driver genes (10). However, some driver genes like MYC in ovarian cancer (11) are not druggable directly, suggesting the application of network-based approaches for identifying druggable interaction partners in the disease module, a local neighborhood of genes linked to the disease mechanism. Moreover, it has been shown that network-based drug repurposing makes predictions that align with the outcome of current clinical drug studies and enables accelerated drug development in cancer (12).

Network-based drug repurposing approaches have proven successful for individual cases but in oncology, these are only accessible to researchers with advanced programming skills (13,14). To make these methods available to a broader scientific community, we present the Cancer Driver Drug Interaction Explorer (CADDIE). CADDIE generates hypotheses for drug development and repurposing in cancer with a graphical user interface that thus not only bioinformaticians or researchers with a computer science background can use this platform but also clinicians and biologists who are looking for candidates for clinical trials or cell line studies. It offers comprehensive gene and protein information collected from different resources as well as an interactive network visualization to inform the selection of seed genes or proteins, i.e. putative drug targets that have been implicated in a cancer type or in an individual via genetic screening and serve as a starting point for network-based drug target drug repurposing candidate discovery. The human interactome can be queried using a number of algorithms and user-defined targets, such that each individual case can be investigated. Users can enter a list of seed genes or proteins and choose among seven different algorithms to prioritize 16761 genes and 6840 drugs.

Comparison with existing tools

A number of in silico tools have been developed for drug repurposing. ACID (15) creates predictions based on individual 3D structures of molecules but can not work with a list of genes or proteins, does not consider second-order interactions in a network approach, or visualize the interactions to enhance the interpretability of the outcome. Gene2Drug (16) predicts gene-drug interactions based on pathway information, but fails to produce results for typical oncogenes such as 'KRAS', 'PTEN' or 'TP53'. Further, it is not possible to search drugs for multiple query genes, and similar to ACID, indirect interactions are not taken into account. NeDRex (10) is a disease-agnostic network medicine platform for disease module identification and drug repurposing. CoVex (9) provides similar functionalities as CADDIE but is limited to SARS-CoV-2. While none of these tools

Table 1. Integrated interaction databases and their incorporated edge types

Name	Data type	Experimental	Literature	Predicted
NCG6	CDG	+	+	+
COSMIC	CDG	+	+	-
IntOGen	CDG	-	-	+ (pipeline)
cancer-genes.org	CDG	-	-	+ (MutPanning)
BioGRID	GGI, DGI	+	+	+
STRING	GGI	+	+	+
APID	GGI	+	-	-
IID	GGI	+	-	+
HTRIdb	GGI	+	-	-
Reactome	DGI	-	+	-
DrugBank	DGI	-	+	-
ChEMBL	DGI	+	+	-
DGIdb	DGI	+	+	-

List of databases implemented in CADDIE. Data types are cancer driver gene (CDG), gene-gene interaction (GGI), and drug-gene interaction (DGI). GGIs encompass interactions of corresponding proteins.

is tailored towards research in oncology, CADDIE applies its network medicine algorithms to ten human gene-gene and drug-gene interaction databases with annotated cancer driver genes, cancer-type specific mutation frequencies, gene expression information, genetically related diseases, and approved anticancer drugs. Thus, in comparison to NeDRex and CoVex, CADDIE allows the user to select the datasets of interest for transparency, reproducibility and comparability. Also, usage of drugs in cancer and related clinical trials is highlighted by integrated datasets, facilitating interpretation of the results in the cancer context The comprehensive visualization in CADDIE allows for intuitive explorative analysis and highlights the interplay between interacting molecules. The network algorithms traverse the complete human interactome searching for directly as well as indirectly operating drugs with the option to consider integrated omics data as edge weights.

MATERIALS AND METHODS

Database

The CADDIE database integrates six gene–gene interaction (GGI) and four drug-gene interaction (DGI) datasets (see Supplementary Material for details). Each dataset provides interactions collected from different sources, e.g. peer-review, text-mining or based on *in silico* predictions (see Table 1).

CADDIE integrates the four cancer driver databases COSMIC (17), NCG6 (18), IntOGen (19) and cancer-gene.org (20) (see Supplementary Figure S1) and mutation frequency data from TCGA to assist users in seed selection. Information on drugs that are classified as 'antineoplastic or immunomodulating' according to the Anatomical Therapeutic Chemical (ATC) classification system was downloaded from the WHO (https://www.whocc.no/atc_ddd_index/) in November 2020. We further created a curated list of approved cancer drugs (CanceRx), which was extracted from the National Cancer Institute website (https://www.cancer.gov/about-cancer/treatment/drugs) and Cancer Research UK website (https://www.cancerresearchuk.org/about-cancer/cancer-in-general/treatment/cancer-drugs/drugs) between 17–26

Table 2. Algorithms for drug target identification and repurposing

Name	Drug target prioritization	Drug repurposing
TrustRank	X	X
Degree Centrality	X	X
Harmonic Centrality	X	X
Betweenness Centrality	X	
KeyPathwayMiner	X	
Multi-level Steiner Tree	X	
Network Proximity		X

Listed are the seven integrated network algorithms used for drug target prioritization and drug repurposing with their respective application cases.

January 2022. Drugs lacking a direct known protein target such as antimetabolites, alkylating agents together with palliative drugs, e.g., those used to treat anemia or chemotherapy induced nausea and vomiting, were not considered. This allows users to distinguish drugs already used as standard of care (SOC) therapy from newly suggested repurposable drugs. Drugs from the Cancer Therapeutics Response Portal (https://portals.broadinstitute.org/ctrp/, version 2) are also implemented in the CADDIE database (21). The drug search can be limited to these drugs, such that all returned drugs were evaluated in the context of cell line studies.

Seed selection

Seed genes are the starting nodes for all implemented network algorithms and thus crucial for the quality of the results. Users can construct a set of seed genes for a selected cancer type based on integrated mutation and tissue-specific gene expression information (see Supplementary Figure S4). Genes with more frequent cancer mutations and high expression are prime drug target candidates (22). Seed genes also can directly be extracted from variant calling files using PolyPhen-2 (23). Lastly, seeds can be selected based on related diseases or uploaded as a list.

Algorithms

CADDIE prioritizes drugs and targets based on their network properties and integrates seven network-based algorithms for both drug target identification and drug repurposing/prioritization (see Table 2) which were previously suggested for drug repurposing in SARS-CoV2 (9). Six different algorithms identify putative drug targets most central to the seed genes in the GGI network and thus indicative of functional relationships (see Figure 1). Four algorithms determine drug candidates in proximity to the seed genes in the DGI network, hence drugs that may have indirect influence on the seed genes. As many putative cancer driver genes are difficult to be targeted directly (11), users may consider running a drug target search prior to a drug search.

Multi-level Steiner Tree (24), KeyPathwayMiner (25) and betweenness centrality can be considered as aggregation approaches as they connect the seeds by finding nodes in between them. The other algorithms are propagation methods which traverse deeper into the network starting from the seed genes. Only the propagating algorithms can identify drug candidates as the drugs lie in the network outside of the disease module. In comparison to the other drug search algorithms, Network Proximity (26) is the most explorative algorithm by design, propagating deeper into the network to propose drugs that are linked to the neighborhood of the seeds.

Due to efficient preprocessing of the networks with graph-tool (https://graph-tool.skewed.de/, version 2.32), the complete human gene-drug interactome can be searched for the optimal results in real-time. The graph-tool module makes it possible to benefit from the speed advantages of C++ while still having the simplicity of python code.

The GGI and DGI networks were augmented with the expression and mutation data stored in the CADDIE database. Using this data as edge weights, the network algorithms can account for the differences in gene expression and somatic mutation frequencies across tumor types. CADDIE further allows users to restrict DGIs to activators or inhibitors. The choice of algorithm and its parameters remains up to the user. It is often beneficial to run multiple analysis tasks with varying settings to cover a broader range of solutions. For this purpose, a task summary function combines the results into one single network while node counts give an intuitive understanding of the importance.

Programmatic access

The CADDIE backend with a task server and the database offers an application programming interface (API) to exchange data with other services. A python package was created to provide a programmatic interface to CADDIE's drug repurposing functionalities. It allows running a large number of tasks using different algorithms and parameters in custom workflows for sophisticated analyses as part of other tools. CADDIE is further available as docker container which encapsulates all dependencies and can be deployed locally to overcome privacy concerns w.r.t. sensitive genomic data (27).

RESULTS

CADDIE web interface

CADDIE is an open access online platform for prioritizing (repurposing) putative drug candidates or targets in cancer using network medicine methods on the combined human gene-gene and drug-gene interactome. Results are visualized online and can be downloaded for further processing. The web interface is freely available and no login is required.

The main feature is the Explorer (see Supplementary Figure S2), which allows users to visualise the interactome of genes and consider known cancer driver genes to find seeds for drug target identification and drug prioritization. Analyses can be triggered with default values or with customized parameters for fine-tuning. Since the full GGI network with 16761 genes would clutter the view, CADDIE shows only cancer type-related genes connected through a minimum spanning tree (see Supplementary Figure S3) (28).
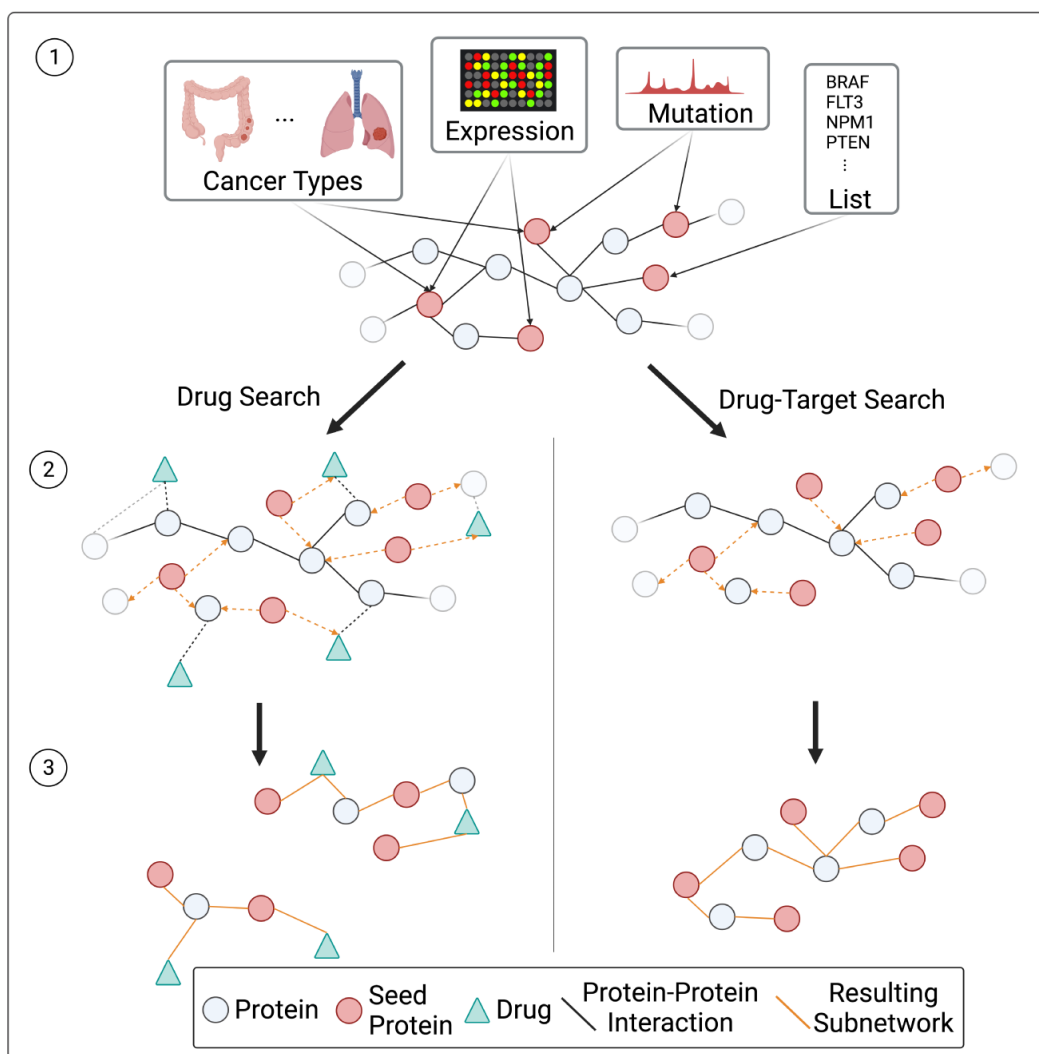
Figure 1. Seed genes for drug target and drug search can be selected from integrated cancer driver databases, via mutation or expression data from TCGA and GTEx using user-defined thresholds, or by uploading a list of genes (1). Based on the seed genes, putative drugs or drug target candidates are identified using a broad selection of network medicine algorithms that traverse the human gene-drug interactome (2). While the drug search reports drugs in proximity to the seed genes, the drug target search returns genes interacting with the disease module spanned by the seed genes (3).

The GGI network visualizes the interactions of the cancer driver genes, helping the user to define a selection of seed nodes that will then be considered for the drug target or drug search. CADDIE also displays complementary information about mutation rates, gene expression, comorbidities, and related cancer types (see Supplementary Figure S4).

Users can choose among several network medicine algorithms (see Supplemental Material for details) for identifying direct or proximal drug targets and related drug (repurposing) candidates. The analysis view shows the results, i.e. genes ranked by their importance according to a user-selected network centrality measure. Detailed information about drugs is highlighted. CADDIE also contrasts importance scores with the node degrees in the interactome (see Supplementary Figure S5 and Supplementary Table S1). This is due to the fact that nodes with a high degree are more likely to be selected by chance, whereas nodes with a low degree and a high score could be more specific to the disease

and thus represent attractive targets. All used parameters are displayed for reproducibility. Task results can be shared by copying the URL and all data can be downloaded.

Showcase sarcoma

Ohshima *et al.* identified cancer driver genes in 16 cancer types based on their frequent amplification and overexpression (29). Their study highlighted sarcoma where AXL was reported as drug target. Based on this finding, we selected the identified genes DNMT3A, FGFR3, ALK, EZH2, HIST1H3B, MDM2, SKP2, MYC, AXL and FLCN as seed genes in CADDIE. We further chose BioGRID (30) as GGI and DGI dataset and selected the sarcoma cancer type from COSMIC to highlight known drivers for this cancer type. Next, we used the TrustRank algorithm (limited to the top 15 drugs with the highest scores, default parameters otherwise) where we included indirect and unapproved drugs for an exhaustive result (see Figure 2).
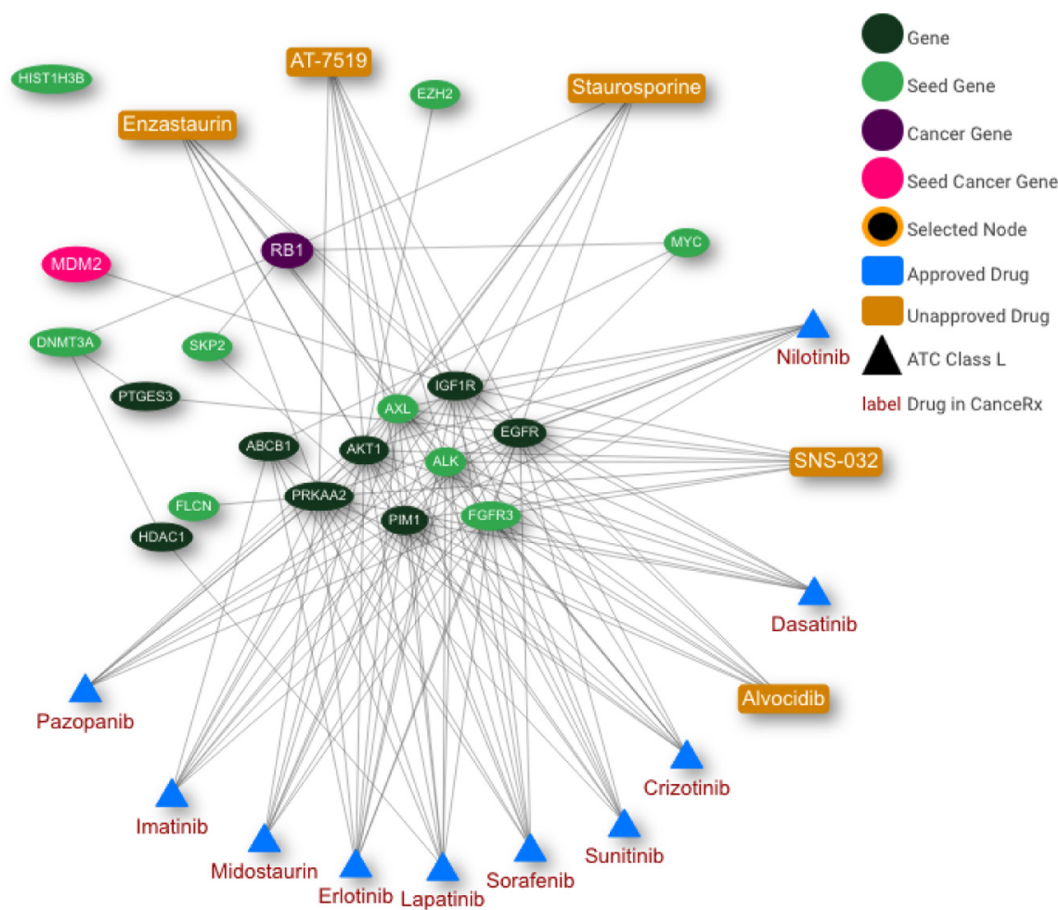
Figure 2. Drug network computed with TrustRank for sarcoma with BioGRID gene–gene and drug–gene interactions. The cancer driver gene information is taken from COSMIC.

Table 3. Drug search results for sarcoma

Name	Approved	ATC L	CanceRx	Score	Degree
Imatinib	yes	yes	yes	1	413
Staurosporine	no	no	no	0.995	393
Sorafenib	yes	yes	yes	0.991	401
Lapatinib	yes	yes	yes	0.984	387
Sunitinib	yes	yes	yes	0.983	388
Alvocidib	no	no	no	0.98	381
Dasatinib	yes	yes	yes	0.978	381
Nilotinib	yes	yes	yes	0.978	386
AT-7519	no	no	no	0.977	377
SNS-032	no	no	no	0.976	376
Erlotinib	yes	yes	yes	0.976	381
Pazopanib	yes	yes	yes	0.975	375
Crizotinib	yes	yes	yes	0.974	377
Midostaurin	yes	yes	yes	0.974	377
Enzastaurin	no	no	no	0.974	376

Listed are the top 15 drug results reported by CADDIE for sarcoma. Each row contains the drug name, approval by FDA, EMA or HC, whether it is listed as antineoplastic or immunomodulating agent (ATC class L) by the WHO, whether it is contained in CanceRx, the normalized score and the node degree in the drug–gene-interactome (BioGRID).

In the resulting network, we notice that the seed genes AXL, ALK and FGFR3 are grouped in the center of the graph as they have interactions with all of the reported drugs. RB1, a gene classified as a driver in sarcoma by COS-MIC (31), is also well connected in our solution, linking protein kinase C inhibitor staurosporine (32) to the seed genes MYC, SKP2 and DNMT3A.

The highest scored drugs imatinib (see Table 3) as well as sorafenib, sunitinib, dasatinib, nilotinib, pazopanib and crizotinib have been studied in sarcoma and are partially approved by the FDA for this use case (33). Further, CADDIE assigned the second highest score to staurosporine, an investigational drug with apoptosis-inducing abilities and synthetic analogs as anticancer drugs in clinical trials (32,34), suggesting it as a putative drug for further research.

CDK-9 is a previously reported drug target in sarcoma (35). The CDK-9 inhibitors alvocidib and SNS-032 have shown effective in treating Ewing sarcoma cells (36,37). Interestingly, the drug AT-7519, also a CDK-9 inhibitor, has not been considered in sarcoma while it is already undergoing clinical trials as an anticancer drug in different cancer types (38).

The approved drug lapatinib has shown modest effectiveness in a phase 2 study in a sarcoma subtype (39). Another approved drug is erlotinib, which has shown significant growth delay in a sarcoma xenograft model (40). Approved drugs represent valuable repurposing candidates as their safety is already tested.

CONCLUSION

For the future, we plan to extend CADDIE with additional network analysis methods (41) and data sources, e.g. oncology data such as cell line study information from PharmacoDB (42) to improve interpretability to the meaning of the predictions. Antineoplastic agents have a broad spectrum of targets and the cause of the intended effect often remains unclear. Databases reporting on cancer driver genes label cancer types differently and show differences in granularity for cancer subtypes. While this lack of harmonization poses a challenge for the users of CADDIE, we decided against limiting the search of seed genes to broadly accepted labels. In summary, CADDIE is the first drug repurposing platform tailored towards oncology. It offers access to a broad set of resources and allows biomedical researchers to suggest and prioritize drug targets and to identify suitable drug repurposing candidates using state-of-the-art network medicine algorithms.

DATA AVAILABILITY

The authors declare that all data supporting the findings of this study are available publicly and their integration is described accordingly within the paper and its supplementary information file. The CADDIE code is available on GitHub (https://github.com/biomedbigdata) in the repositories CADDIE-frontend, CADDIE-backend and caddiepy.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Meacham,C.E. and Morrison,S.J. (2013) Tumour heterogeneity and cancer cell plasticity. *Nature*, **501**, 328–337.
2. Nogales,C., Mamdouh,Z.M., List,M., Kiel,C., Casas,A.I. and Schmidt,H.H. (2022) Network pharmacology: curing causal mechanisms instead of treating symptoms. *Trends Pharmacol. Sci.*, **43**, 136–150.
3. Zhang,Z., Zhou,L., Xie,N., Nice,E.C., Zhang,T., Cui,Y. and Huang,C. (2020) Overcoming cancer therapeutic bottleneck by drug repurposing. *Signal Transduct. Targeted Ther.*, **5**, 113.
4. Verma,M. (2012) Personalized Medicine and Cancer. *J. Pers, Med.*, **2**, 1–14.
5. Rubin,E.H. and Gilliland,D.G. (2012) Drug development and clinical trials—the path to an approved cancer drug. *Nat. Rev. Clin. Oncol.*, **9**, 215–222.
6. Pushpakom,S., Iorio,F., Eyers,P.A., Escott,K.J., Hopper,S., Wells,A., Doig,A., Guilliams,T., Latimer,J., McNamee,C. *et al.* (2018) Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug Disc.*, **18**, 41–58.
7. Cheng,F., Desai,R.J., Handy,D.E., Wang,R., Schneeweiss,S., Barabási,A.L. and Loscalzo,J. (2018) Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.*, **9**, 2691.
8. Zhou,Y., Hou,Y., Shen,J., Huang,Y., Martin,W. and Cheng,F. (2020) Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Disc.*, **6**, 14.
9. Sadegh,S., Matschinske,J., Blumenthal,D.B., Galindez,G., Kacprowski,T., List,M., Nasirigerdeh,R., Oubounyt,M., Pichlmair,A., Rose,T.D. *et al.* (2020) Exploring the SARS-CoV-2 virus-host-drug interactome for drug repurposing. *Nat. Commun.*, **11**, 3518.
10. Sadegh,S., Skelton,J., Anastasi,E., Bernett,J., Blumenthal,D.B., Galindez,G., Salgado-Albarrán,M., Lazareva,O., Flanagan,K., Cockell,S. *et al.* (2021) Network medicine for disease module identification and drug repurposing with the NeDRex platform. *Nat. Commun.*, **12**, 6848.
11. Zeng,M., Kwiatkowski,N.P., Zhang,T., Nabet,B., Xu,M., Liang,Y., Quan,C., Wang,J., Hao,M., Palakurthi,S. *et al.* (2018) Targeting MYC dependency in ovarian cancer through inhibition of CDK7 and CDK12/13. *eLife*, **7**. e39030.
12. Cheng,F., Lu,W., Liu,C., Fang,J., Hou,Y., Handy,D.E., Wang,R., Zhao,Y., Yang,Y., Huang,J. *et al.* (2019) A genome-wide positioning systems network algorithm for in silico drug repurposing. *Nat. Commun.*, **10**, 3476.
13. Fahimian,G., Zahiri,J., Arab,S.S. and Sajedi,R.H. (2020) RepCOOL: Computational drug repositioning via integrating heterogeneous biological networks. *J. Trans. Med.*, **18**, 375.
14. Gottlieb,A., Stein,G.Y., Ruppin,E. and Sharan,R. (2011) PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.*, **7**, 496.
15. Wang,F., Wu,F.X., Li,C.Z., Jia,C.Y., Su,S.W., Hao,G.F. and Yang,G.F. (2019) ACID: a free tool for drug repurposing using consensus inverse docking strategy. *J. Cheminformatics*, **11**, 73.
16. Napolitano,F., Carrella,D., Mandriani,B., Pisonero-Vaquero,S., Sirci,F., Medina,D.L., Brunetti-Pierri,N. and Bernardo,D.D. (2018) gene2drug: a computational tool for pathway-based rational drug repositioning. *Bioinformatics*, **34**, 1498–1505.
17. Tate,J.G., Bamford,S., Jubb,H.C., Sondka,Z., Beare,D.M., Bindal,N., Boutselakis,H., Cole,C.G., Creatore,C. *et al.* (2019) COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.*, **47**, D941–D947.
18. Repana,D., Nulsen,J., Dressler,L., Bortolomeazzi,M., Venkata,S.K., Tourna,A., Yakovleva,A., Palmieri,T. and Ciccarelli,F.D. (2019) The Network of Cancer Genes (NCG): A comprehensive catalogue of known and candidate cancer genes from cancer sequencing screens 06 Biological Sciences 0604 Genetics 11 Medical and Health Sciences 1112 Oncology and Carcinogenesis 06 Biological Sciences 0601 Biochemistry and Cell Biology. *Genome Biol.*, **20**, 1.
19. Martínez-Jiménez,F., Muiños,F., Sentís,I., Deu-Pons,J., Reyes-Salazar,I., Arnedo-Pac,C., Mularoni,L., Pich,O., Bonet,J., Kranas,H. *et al.* (2020) A compendium of mutational cancer driver genes. *Nat. Rev. Cancer*, **20**, 555–572.
20. Dietlein,F., Weghorn,D., Taylor-Weiner,A., Richters,A., Reardon,B., Liu,D., Lander,E.S., Allen,E.M.V. and Sunyaev,S.R. (2020) Identification of cancer driver genes based on nucleotide context. *Nat. Genet.*, **52**, 208–218.
21. Rees,M.G., Seashore-Ludlow,B., Cheah,J.H., Adams,D.J., Price,E.V., Gill,S., Javaid,S., Coletti,M.E., Jones,V.L., Bodycombe,N.E. *et al.*

(2016) Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat. Chem. Biol.*, **12**, 109–116.

22. Benson,J.D., Chen,Y.N.P., Cornell-Kennon,S.A., Dorsch,M., Kim,S., Leszczyniecka,M., Sellers,W.R. and Lengauer,C. (2006) Validating cancer drug targets. *Nature*, **441**, 451–456.

23. Adzhubei,I.A., Schmidt,S., Peshkin,L., Ramensky,V.E., Gerasimova,A., Bork,P., Kondrashov,A.S. and Sunyaev,S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods*, **7**, 248–249.

24. Graham,R.L. and Hell,P. (1985) On the History of the Minimum Spanning Tree Problem. *Ann. Hist. Comput.*, **7**, 43–57.

25. Alcaraz,N., List,M., Dissing-Hansen,M., Rehmsmeier,M., Tan,Q., Mollenhauer,J., Ditzel,H.J. and Baumbach,J. (2016) Robust de novo pathway enrichment with KeyPathwayMiner 5 [version 1; referees: 2 approved]. *F1000Research*, **5**, 1531.

26. Guney,E., Menche,J., Vidal,M. and Barábasi,A.L. (2016) Network-based in silico drug efficacy screening. *Nat. Commun.*, **7**, 10331.

27. Naveed,M., Ayday,E., Clayton,E.W., Fellay,J., Gunter,C.A., Hubaux,J.P., Malin,B.A. and Wang,X. (2015) Privacy in the Genomic Era. *ACM Comput. Surv. (CSUR)*, **48**, 6.

28. Ahmed,R., Angelini,P., Sahneh,F.D., Efrat,A., Glickenstein,D., Gronemann,M., Heinsohn,N., Kobourov,S.G., Spence,R., Watkins,J. *et al.* (2019) Multi-level Steiner Trees. *J. Exp. Algorith. (JEA)*, **24**, 1–22.

29. Ohshima,K., Hatakeyama,K., Nagashima,T., Watanabe,Y., Kanto,K., Doi,Y., Ide,T., Shimoda,Y., Tanabe,T., Ohnami,S. *et al.* (2017) Integrated analysis of gene expression and copy number identified potential cancer driver genes with amplification-dependent overexpression in 1,454 solid tumors. *Scientific Rep.*, **7**, 641.

30. Oughtred,R., Rust,J., Chang,C., Breitkreutz,B.J., Stark,C., Willems,A., Boucher,L., Leung,G., Kolas,N., Zhang,F. *et al.* (2021) The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.*, **30**, 187–200.

31. Liu,Y., Sánchez-Tilló,E., Lu,X., Clem,B., Telang,S., Jenson,A.B., Cuatrecasas,M., Chesney,J., Postigo,A. and Dean,D.C. (2013) Rb1 family mutation is sufficient for sarcoma initiation. *Nat. Commun.*, **4**, 2650.

32. Wang,Y.F., Hsieh,Y.F., Lin,C.L., Lin,J.L., Chen,C.Y., Chiou,Y.H. and Chou,M.C. (2004) Staurosporine-induced G2/M arrest in primary effusion lymphoma BCBL-1 cells. *Ann. Hematol.*, **83**, 739–744.

33. Forscher,C., Mita,M. and Figlin,R. (2014) Targeted therapy for sarcomas. *Biol. Tar. Ther.*, **8**, 91.

34. Malsy,M., Bitzinger,D., Graf,B. and Bundscherer,A. (2019) Staurosporine induces apoptosis in pancreatic carcinoma cells PaTu 8988t and Panc-1 via the intrinsic signaling pathway. *Eur. J. Med. Res.*, **24**, 5.

35. Cassandri,M., Fioravanti,R., Pomella,S., Valente,S., Rotili,D., Baldo,G.D., Angelis,B.D., Rota,R. and Mai,A. (2020) CDK9 as a Valuable Target in Cancer: From Natural Compounds Inhibitors to Current Treatment in Pediatric Soft Tissue Sarcomas. *Front. Pharmacol.*, **11**, 1230.

36. Flores,G. and Grohar,P.J. (2021) One oncogene, several vulnerabilities: EWS/FLI targeted therapies for Ewing sarcoma. *J. Bone Oncol.*, **31**, 100404.

37. Flores,G., Everett,J., Osgood,C., Madaj,Z. and Grohar,P.J. (2017) CDK 7/9 inhibition amplifies mithramycin's suppression of Ewing sarcoma cell proliferation. *FASEB J.*, **31**, 178.10.

38. Mandal,R., Becker,S. and Strebhardt,K. (2021) Targeting CDK9 for anti-cancer therapeutics. *Cancers*, **13**, 2181.

39. Stacchiotti,S., Tamborini,E., Vullo,S.L., Bozzi,F., Messina,A., Morosi,C., Casale,A., Crippa,F., Conca,E., Negri,T. *et al.* (2013) Phase II study on lapatinib in advanced EGFR-positive chordoma. *Ann. Oncol.*, **24**, 1931–1936.

40. Bandyopadhyay,A., Favours,E., Phelps,D.A., Pozo,V.D., Ghilu,S., Kurmashev,D., Michalek,J., Trevino,A., Guttridge,D., London,C. *et al.* (2018) Evaluation of patritumab with or without erlotinib in combination with standard cytotoxic agents against pediatric sarcoma xenograft models. *Pediatr. Blood Cancer*, **65**, e26870.

41. Bernett,J., Krupke,D., Sadegh,S., Baumbach,J., Fekete,S.P., Kacprowski,T., List,M. and Blumenthal,D.B. (2022) Robust disease module mining via enumeration of diverse prize-collecting Steiner trees. *Bioinformatics*, **38**, 1600–1606.

42. Smirnov,P., Kofia,V., Maru,A., Freeman,M., Ho,C., El-Hachem,N., Adam,G.A., Ba-Alawi,W., Safikhani,Z. and Haibe-Kains,B. (2018) PharmacoDB: an integrative database for mining in vitro anticancer drug screening studies. *Nucleic Acids Res.*, **46**, D994–D1002.