NAR Cancer, 2025, **7**, zcae047
https://doi.org/10.1093/narcan/zcae047
Advance access publication date: 15 January 2025
Cancer Data Resource

The Atlas of Protein–Protein Interactions in Cancer (APPIC)—a webtool to visualize and analyze cancer subtypes

Benjamin Ahn[1,†], Charissa Chou[2,†], Caden Chou[1], Jennifer Chen[1], Amelia Zug[1], Yigit Baykara[1,2], Jessica Claus[1,2], Sean M. Hacking[3], Alper Uzun[1,4,5,6,*] and Ece D. Gamsiz Uzun[1,2,4,5,6,*]

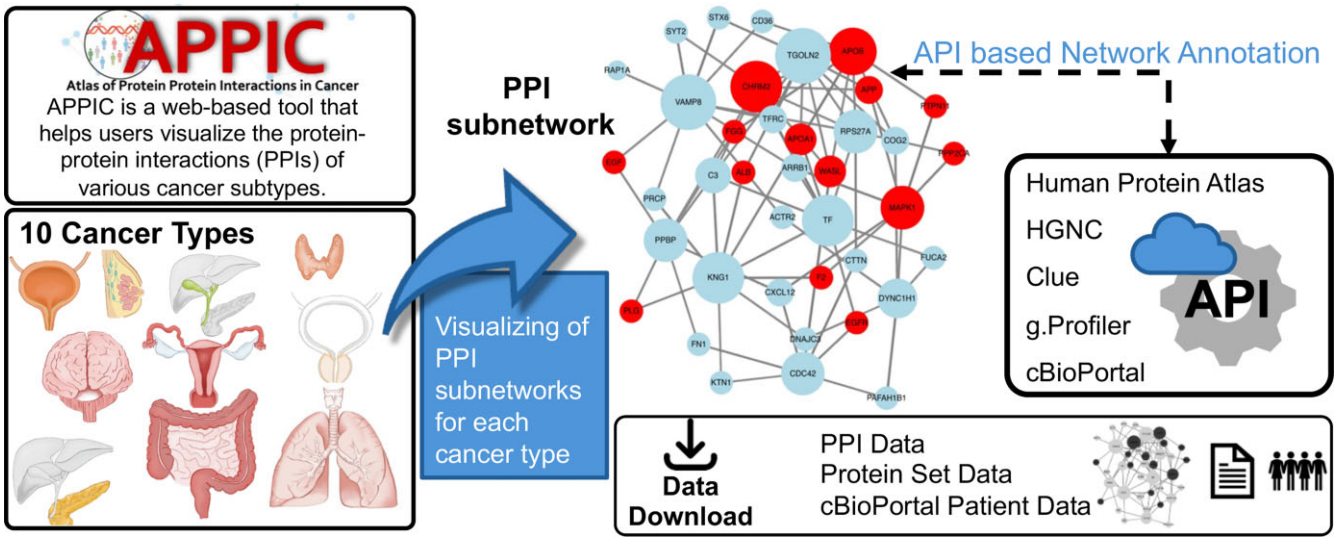
[1]Department of Pathology and Laboratory Medicine, The Warren Alpert Medical School of Brown University, 593 Eddy Street, Providence, RI 02903, USA
[2]Department of Pathology and Laboratory Medicine, Rhode Island Hospital, 593 Eddy Street, Providence, RI 02903, USA
[3]Department of Pathology, NYU Grossman School of Medicine, 550 1st Ave., New York, NY 10016, USA
[4]Legoretta Cancer Center, Brown University, 70 Ship Street, Providence, RI 02903, USA
[5]Brown Center for Clinical Cancer Informatics and Data Science (CCIDS), Brown University, 593 Eddy Street, Providence, RI 02903, USA
[6]Center for Computational Molecular Biology, Brown University, 164 Angell Street, Providence, RI 02906, USA

[*]To whom correspondence should be addressed. Tel: +1 401 444 4000; Fax: +1 401 444 3109; Email: dilber_gamsiz@brown.edu
Correspondence may also be addressed to Alper Uzun. Tel: +1 401 863 1000; Fax: +1 401 444 3109; Email: alper_uzun@brown.edu
[†]The first two authors should be regarded as Joint First Authors.

Abstract

Cancer is a complex disease with heterogeneous mutational and gene expression patterns. Subgroups of patients who share a phenotype might share a specific genetic architecture including protein–protein interactions (PPIs). We developed the Atlas of Protein–Protein Interactions in Cancer (APPIC), an interactive webtool that provides PPI subnetworks of 10 cancer types and their subtypes shared by cohorts of patients. To achieve this, we analyzed publicly available RNA sequencing data from patients and identified PPIs specific to 26 distinct cancer subtypes. APPIC compiles biological and clinical information from various databases, including the Human Protein Atlas, Hugo Gene Nomenclature Committee, g:Profiler, cBioPortal and Clue.io. The user-friendly interface allows for both 2D and 3D PPI network visualizations, enhancing the usability and interpretability of complex data. For advanced users seeking greater customization, APPIC conveniently provides all output files for further analysis and visualization on other platforms or tools. By offering comprehensive insights into PPIs and their role in cancer, APPIC aims to support the discovery of tumor subtype-specific novel targeted therapeutics and drug repurposing. APPIC is freely available at https://appic.brown.edu.

Graphical abstract



Received: August 21, 2024. **Revised:** November 27, 2024. **Editorial Decision:** December 22, 2024. **Accepted:** January 3, 2025

Introduction

In 2024, around 2 million new cancer cases are expected to be diagnosed and 611 720 cancer deaths are projected to occur in the United States (1). According to Centers for Disease Control and Prevention (CDC), cancer is the second leading cause of death in the United States, exceeded only by heart disease (2). Cancer is a complex disease that is often attributed to perturbations in gene networks (3,4). Despite numerous advances in medical research, our understanding of the intricate molecular and genetic interactions that drive cancer formation, progression and metastasis is still evolving (5). Given the heterogeneous nature of cancer, which can vary greatly between individuals and even within a single tumor (6,7), there is an urgent need for tools that can decode these complex interactions. Advances in high-throughput sequencing technologies, genome-wide association studies and bioinformatics methods have improved our understanding in the genetics of complex diseases, including cancer. In this respect, network biology-based approaches have emerged as powerful tools as they have potential to reveal the biological mechanisms of complex diseases such as cancer by bridging phenotype–genotype information (8,9). These large-scale protein–protein interaction (PPI) maps indicate that genes linked to similar phenotypes often physically interact at the protein level (10). Furthermore, similar phenotypes tend to occupy neighboring network spaces, a phenomenon observed in both model organisms and humans. The availability of several publicly accessible PPI databases marks a significant advancement in this domain (9,11). PPI databases are critical resources for understanding the complex networks that underpin cellular functions and disease mechanisms. STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) stands out for its integration of data from experimental results, computational predictions and public text collections (11). Additionally, databases such as BioGRID, IntAct, the Human Protein Reference Database (HPRD), MINT, ConsensusPathDB, FunCoup, GeneMANIA, HumanBase, HumanNet and IID provide extensive PPI data for various organisms and pathways (9,12–20). Recognizing the potential utility of PPI network analysis in complex disease research, in our previous study we developed Proteinarium, a specialized tool designed to identify patient clusters with shared PPI networks (21). Proteinarium integrates multi-sample PPI analysis and visualization capabilities, allowing researchers to examine PPI networks that drive different disease phenotypes. Proteinarium stands out for its ability to construct patient-specific PPI networks by integrating gene expression data with experimentally validated PPI information. It identifies clusters of patients with shared network similarities, offering a unique view into molecular interactions that may drive distinct disease phenotypes. By highlighting hub proteins with therapeutic potential, Proteinarium supports the identification of novel drug targets, making it an essential tool in precision medicine. Leveraging STRING's experimentally validated PPI data, Proteinarium facilitates detailed exploration of PPI networks specific to each patient, aiding in the understanding of molecular mechanisms underlying various tumor types and identifying potential novel therapeutic targets (22).

Using Proteinarium, we identified the PPI networks specific to 26 cancer subtypes across 10 tissue types, including bladder, brain, breast, colon/colorectal, gallbladder, lung, ovarian, pancreas, prostate and thyroid. By integrating the PPI networks identified into an interactive user interface, we developed the Atlas of Protein–Protein Interactions in Cancer (APPIC), a novel tool in cancer research. APPIC is an interactive web application that visualizes PPI networks in several tumor types and their subtypes as well as aggregates clinical and biological information from databases such as the Human Protein Atlas (HPA), HUGO Gene Nomenclature Committee (HGNC), g:Profiler, cBioPortal and Clue.io (23–26). We used publicly available RNA-sequencing (RNA-seq) data from patients for the Proteinarium analysis. This web application functions as a dynamic tool to support the development of new therapeutic strategies, aid in drug repurposing efforts and further elucidate mechanisms behind tumor subtype formation.

APPIC

APPIC is an interactive webtool that presents the PPI subnetworks specific to 26 cancer subtypes across 10 tissue types. Users can select a tissue type or subtype to explore the PPI subnetworks. Based on the selection, APPIC generates the 2D or 3D PPI network, and allows the users to browse the information aggregated from databases such as HPA, HGNC, Clue.io and g:Profiler. The size of a node representing a protein is proportional to the number of connections with other proteins in the PPI. Proteins with the highest number of connections, known as 'hub proteins', are potential drug targets. Studies showed that hub proteins in PPI networks are preferentially selected as drug targets due to their high connectivity and ability to propagate effects. However, being a hub protein alone does not ensure therapeutic potential without further analysis (27,28).

Information on a specific protein in a selected PPI network is obtained from HPA and HGNC. The g:Profiler section displays the biological pathways in the selected PPI network. Clue.io section searches for existing drugs that target the selected protein. The proteins with potential drug targets in the PPI network are highlighted in red. cBioPortal section displays the Kaplan–Meier survival curve for the specific patient cluster used in the analysis of the PPI network. Users can download a list of patient IDs sharing the consensus network and the clinical data from cBioPortal for further analysis (Figure 1). APPIC is freely available at https://appic.brown.edu.

RNA-seq Data Processing and Proteinarium Analysis

RNA-seq data from 10 cancer types, including bladder, brain, breast, colon/colorectal, gallbladder, lung, ovarian, pancreas, prostate and thyroid carcinomas, were downloaded from cBioPortal. The utilized datasets were obtained as messenger RNA expression z-scores that were precalculated relative to other tumor samples within the cohort, allowing the users to normalize expression values across all samples. Pseudogenes were filtered from the datasets. Subsequently, for each patient, we ranked the genes in descending order based on expression levels to compile a list of 'seed genes', which was employed to execute Proteinarium (21). To determine the number of seed genes to be used in Proteinarium analysis, we prepared lists of the top N highly expressed genes ($N = 50$, 100, 150, 200, 250 and 300) for each patient (Figure 2). Genes are ranked for each patient based on their z-scores. The top N highly ranked genes,

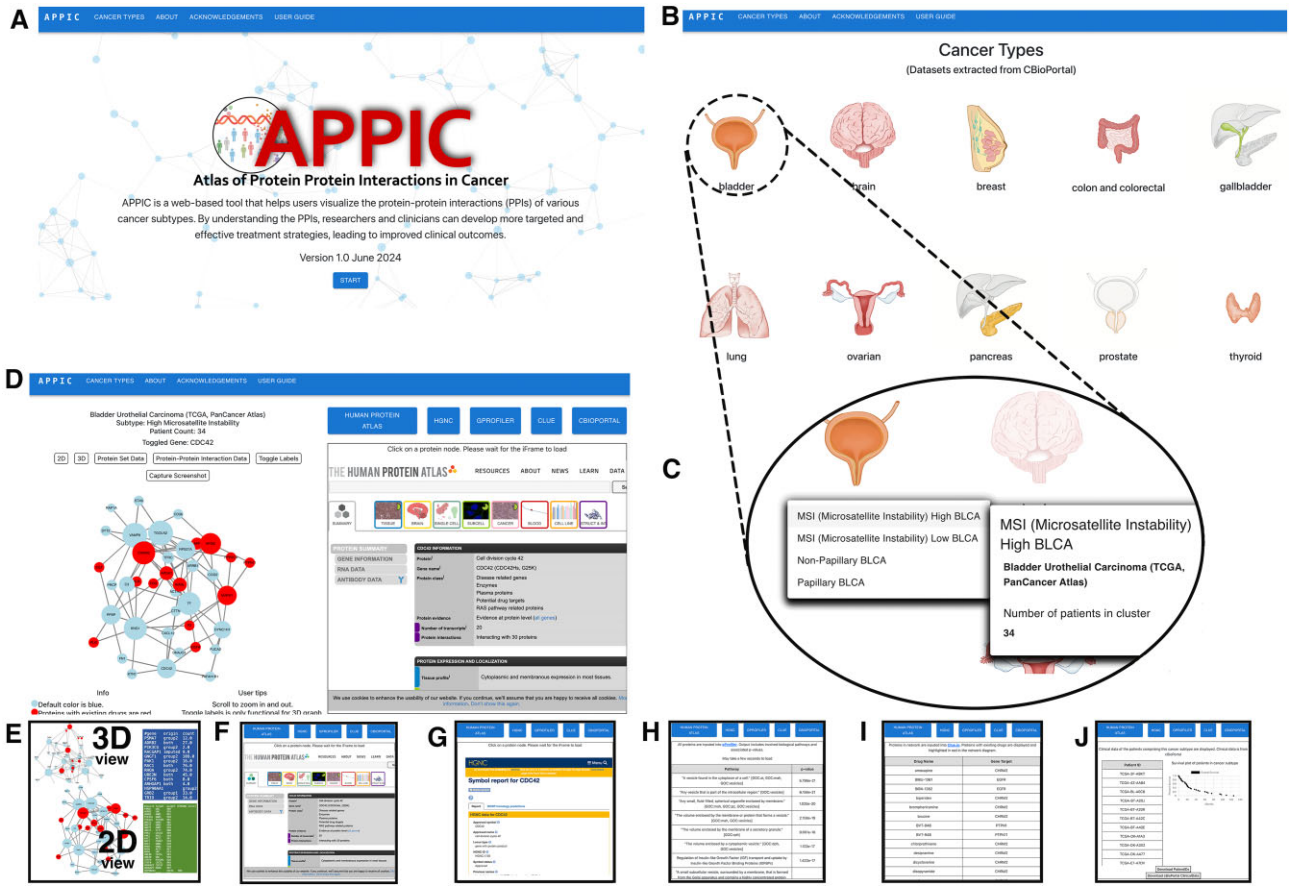


Figure 1. (**A**) APPIC main page: the primary interface of the APPIC platform. (**B**) Interactive Cancer Types page: a page allowing the selection and exploration of various tumor types. (**C**) Example query: display for non-papillary urothelial carcinoma of the bladder (TCGA, PanCancer Atlas) with a cohort of 269 patients. (**D**) Results page for the query: left side shows the interactive consensus network for 269 patients. Nodes (or proteins) are blue by default, and nodes with drugs available for repurposing are highlighted in red. Right side provides detailed information about each node (or protein) from HPA, HGNC, Clue.io, g:Profiler and cBioPortal when selected. (**E**) Visualization options: 2D and 3D PPI networks, downloadable data for PPIs and screen capture for the PPI network. (**F**) Dynamic link to HPA: direct link to the HPA site for the queried protein in the network. (**G**) HGNC information: details about the protein as provided by HGNC. (**H**) Network annotation: biological pathways and associated *P*-values precalculated by g:Profiler. (**I**) Proteins with existing drugs (from Clue.io) are highlighted in red within the network diagram and listed in a table. (**J**) Clinical data display: survival plot of patients in the selected tumor subtype, based on clinical data from cBioPortal. Patient IDs and clinical data are available for download for further analysis.

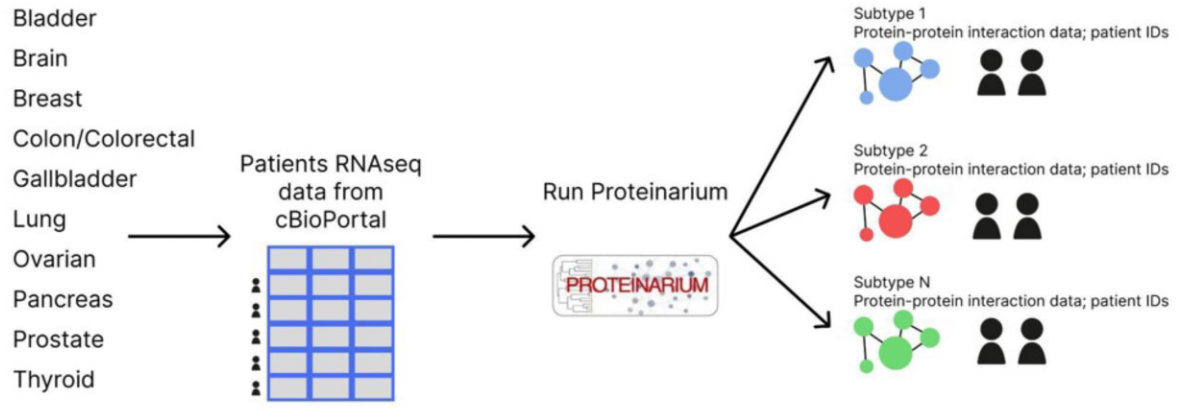


Figure 2. Identifying PPIs specific to cancer subtypes using Proteinarium. Proteinarium used RNA-seq data from patients with 10 tumor types to identify PPI networks specific to 26 cancer subtypes.

as determined by their z-scores, are selected as seed genes for the construction of PPI networks. Proteinarium is run with these top N seed genes. One of the main outputs of Proteinarium is a dendrogram, which clusters patients based on their network similarities. To find the most stable dendrogram with the optimum number of seed genes, the dendrograms are compared using the cophenetic correlation coefficient (CCC) (29). The CCC measures how accurately the dendrogram preserves the pairwise distances between the original data points. The best representative dendrogram is chosen by selecting the one with the highest CCC score and the least number of seed genes. Further details of this method were described in our study by Hacking *et al.* (22).

Proteinarium and PPI Network Construction

We used Proteinarium to build the PPI networks. Proteinarium is a unique tool designed for multi-sample PPI network analysis. It builds patient-specific PPI networks by integrating gene expression data, such as RNA-seq, with experimentally validated PPI information from STRING, allowing for the identification of consensus PPI networks specific to a patient cohort. A key strength is its ability to cluster patients based on network similarities, aiding in understanding disease heterogeneity. Proteinarium also offers layered visualization of networks, providing a clear and customizable view of PPIs. Additionally, its capacity to reveal hub proteins makes it a valuable tool for identifying novel therapeutic targets, particularly in diseases such as cancer. Proteinarium requires a list of genes (seed genes) for each sample as an input and generates a PPI network for each patient by mapping the seed gene list onto experimentally validated PPI information derived from the STRING database (11). Proteinarium identifies clusters of samples based on network similarities and presents the clusters in a dendrogram. It uses Dijkstra's algorithm for the shortest path and the Jaccard index to build a network similarity matrix of PPI between samples (21). In constructing the PPI networks using APPIC, the values for key parameters, maxPathLength and maxPathCost, were carefully chosen to balance network complexity and biological relevance. maxPathLength was set to 2, meaning that a maximum of one intermediary node is allowed between seed proteins. This value was selected to capture relevant indirect interactions while avoiding over-expansion of the network with potentially unrelated nodes. Keeping the path length short ensures that the interactions reflect more direct biological relationships, which is critical for interpreting the network in the context of cancer subtypes. maxPathCost was set to 2000, based on interaction confidence scores from the STRING database (where 1000 is the maximum score for confidence). This threshold was chosen to focus on high-confidence interactions, ensuring that only interactions with scores above 800 are included. This approach helps filter out lower-confidence connections and improves the biological relevance and robustness of the resulting networks. These parameters ensure that APPIC provides users with concise yet biologically meaningful networks that reflect high-confidence interactions. Patients are clustered according to the similarity of their PPI networks. Consensus PPI networks are then formed based on these clusters. The final output includes gene set and gene interaction files for the consensus PPI networks, as well as a list of patient IDs for each tumor subtype (21).

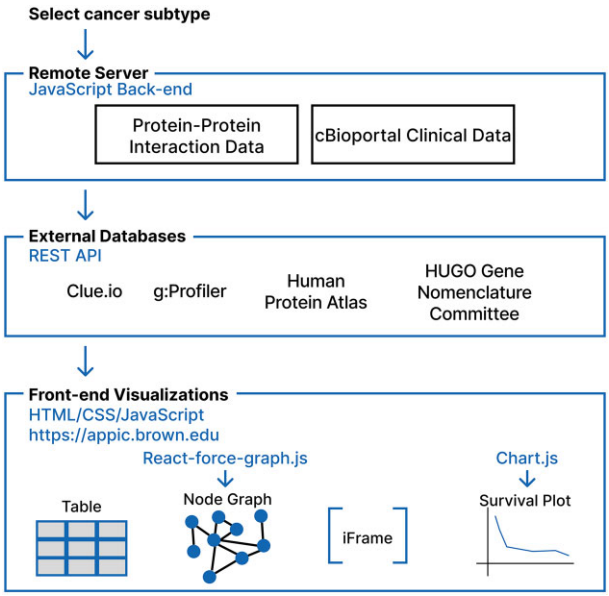


Figure 3. APPIC infrastructure. Users select the cancer type that serves as the initial input. APPIC retrieves and processes the PPI data based on the cancer type, which is stored on the server, to generate node graphs. APPIC extract cBioPortal data to generate survival plots. PPI data are further parsed, organized and sent to g:Profiler and Clue.io using their REST APIs. Responses are received as JSON objects that are displayed as tables on the front end. iFrames are used to embed HTML content of HPA and HGNC into the front end.

System Description

The front end of the webtool was developed using HTML/CSS/JavaScript and the React JavaScript framework (30). The backend was developed using JavaScript (Figure 3). To render the network diagrams, PPI data files are parsed and built using the React-Force-Graph library. To interact with HPA and HGNC, APPIC uses an iFrame to embed the web pages directly into the front end. APIs were used for g:Profiler and Clue.io to perform real-time search of biological pathways and existing drug targets relevant to the protein selected by the user. The information received is presented in tables built in HTML/CSS. cBioPortal datasets were downloaded as CSV files and stored in the backend of the webtool. APPIC parses these files to display the clinical data in HTML/CSS/JavaScript. The webtool including backend, front end and data files is hosted on a Brown University server.

Case Study

This case study aims to demonstrate how APPIC can be used to explore the PPI networks in urothelial carcinoma (bladder cancer) (TCGA, PanCancer Atlas), specifically focusing on the high microsatellite instability (MSI-H) subtype (Figure 4).

Step 1: Identify hub proteins

Within the urothelial carcinoma (MSI-H) PPI network, hub proteins—proteins with the highest number of interactions—can be identified. These proteins are often essential for cancer progression and could serve as valuable targets for drug development or repurposing. The size of each node in the PPI network corresponds to the number of its connections. For instance, the most connected protein in the network is VAMP8,

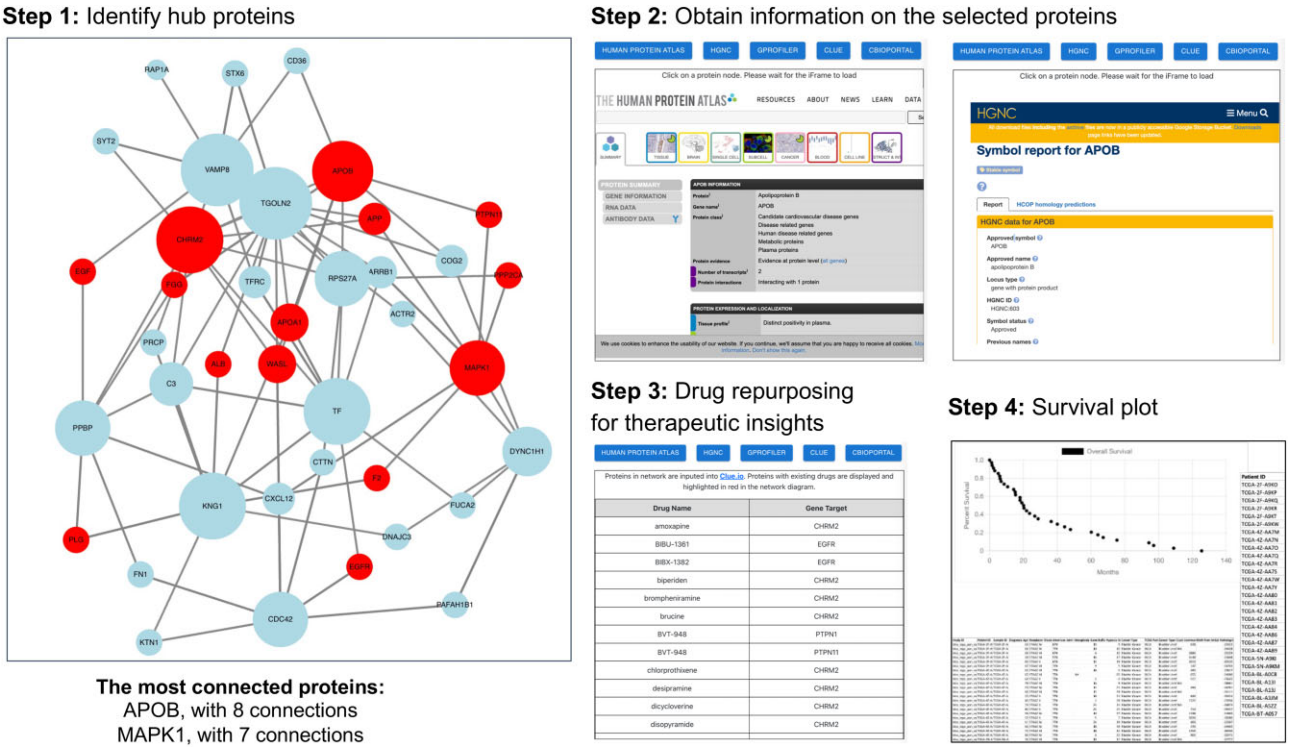


Figure 4. Case study to visualize PPI network specific to MSI-H bladder urothelial carcinoma (TCGA, PanCancer Atlas).

with nine connections. The next two most connected proteins are APOB, with eight connections, and MAPK1, with seven connections.

Step 2: Obtain information on the selected proteins

HPA tab on the panel right side of the network can be accessed by clicking the node on the PPI network and provides information about the selected proteins. Users can also retrieve gene information from the HGNC tab by clicking on a node and selecting the HGNC tab. g:Profiler maps user-provided protein lists to various biological databases and identifies statistically significant enrichments in terms of biological processes, pathways and regulatory motifs. In this example, all proteins in the PPI network are used as input for g:Profiler.

Step 3: Drug repurposing for therapeutic insights

Clue.io was integrated into APPIC to allow the users to identify approved drugs that interact with key proteins in the PPI network. Researchers can identify candidates for further preclinical or clinical testing in the MSI-H urothelial carcinoma subtype. In the PPI network, the nodes, which are colored in red, indicate that drugs are available for those proteins, i.e. CHRM2, APOB, MAPK, WASL, APP, PTPN11, PPP2CA, F2, EGFR, PLG, EGF and ALB in the PPI network shown in Figure 4. The names of these drugs for those proteins can be obtained by clicking on the 'Clue' tab. There are total of 19 potential drugs for MAPK1 and 1 drug for APOB.

Step 4: Survival plot

APPIC provides integrated survival plots from cBioPortal. Survival plots from the patients specifically contributing in a PPI network can be visualized by clicking on 'cBioPortal' tab. This analysis aids in identifying proteins potentially associated with poor survival outcomes (Figure 4).

Discussion

We developed APPIC, an interactive webtool for researchers and clinicians to visualize and analyze the PPI subnetworks specific to 26 cancer subtypes. The integration of databases such as HPA, HGNC, g:Profiler, Clue.io and cBioPortal makes APPIC a unique platform to further study the mechanisms driving cancer subtypes and develop novel precision medicine strategies.

To construct the PPI networks for each patient sample, we chose Proteinarium, a multi-sample PPI network analysis and visualization tool, due to its specialized capabilities in identifying patient clusters with shared PPI networks. Proteinarium excels at building consensus PPI networks in a cohort of patients by using gene expression data from patients to reveal unique molecular mechanisms. Its ability to cluster patients based on network similarities enables the identification of subgroups, aiding in the discovery of potential therapeutic targets. Proteinarium highlights densely connected proteins and identifies potential drug targets. Proteinarium employs Dijkstra's algorithm to construct high-confidence interaction graphs, which is crucial for delineating distinct cancer types or subtypes. It offers robust visualization options, including layered graphs for patient clusters, and supports customizability with user-defined parameters. Designed to handle the RNA-seq data, Proteinarium efficiently supports multi-sample analysis by leveraging the STRING database for comprehensive network construction. By utilizing Proteinarium, we built detailed PPI networks for each cancer subtype, aiding in the discovery of potential novel targeted therapeutics and drug repurposing.

One of the significant features of APPIC is the visualization and analysis of PPIs, as it provides a user-friendly interface enabling a comprehensive understanding of molecular interactions across a diverse range of cancer subtypes. APPIC provides existing drug targets for selected proteins. This is

Table 1. Comparison of cancer PPI tools: APPIC, HMNPPID, PINA 3.0 and OncoPPi

Tool	Data source	Key features	Visualization	Focus
APPIC	RNA-seq data and patient survival data from cBioPortal, HPA, HGNC, g:Profiler, Clue.io, STRING	Visualizes cancer subtype-specific PPI networks, integrates multiple datasets, drug repurposing, survival analysis	2D/3D visualizations of PPI networks with interactive tools	26 subtypes across 10 cancer types (bladder, brain, breast, etc.) with emphasis on drug targets and patient survival
HMNPPID	PubMed abstracts using PPIExtractor, data from various biomedical texts	Provides cancer-specific PPI networks for malignant neoplasms, VisualPPI for network visualization	VisualPPI with multiple layout modes and detailed analysis options	PPIs in malignant neoplasms across 171 cancer types
PINA 3.0	RNA-seq data from TCGA, proteomics data from CPTAC	Integrates tumor-type-specific interactome with RNA-seq and proteomics, mutation drivers and therapeutic targets	Cytoscape.js-based interactive visualization and network customization	Tumor context-specific interactome analysis and cancer driver identification
OncoPPi	Cancer genes from lung cancer cells; PPI data from STRING, BioGrid, IntAct, GeneMANIA, BioPlex	Focuses on cancer-related PPIs and vulnerabilities, interaction discovery, mutual exclusivity analysis	Interactive network visualization with focus on mutual exclusivity and functional associations	Lung cancer PPI landscape expansion, cancer vulnerabilities, therapeutic targets

especially important as the most connected or 'hub' proteins are potential drug targets (27,28). In our previous study using Proteinarium to study the RNA-seq data from a cohort of patients with mismatch repair (MMR) breast cancer, we showed that the PPI network specific to MMR-deficient breast cancer included highly connected clusters of histone proteins. Importantly, the MMR-intact breast cancer-specific PPI network was distinct from the MMR-deficient breast cancer-specific PPI network (22). As histone lysine methyltransferases are potential drug targets in breast cancer, our previous study showed the utility of PPI network analysis including Proteinarium in novel drug target discovery and drug repurposing specific to cancer subtypes. APPIC was developed with the purpose of expanding our efforts in novel drug target discovery and drug repurposing using PPI network analysis to other cancer types and their subtypes. Additionally, network annotations show each protein's function and connect to the HPA for more detailed protein information.

APPIC distinguishes itself from other cancer PPI tools such as HMNPPID, PINA 3.0 and OncoPPi through its comprehensive integration of multiple data sources and its focus on patient-specific PPI networks across various cancer subtypes (31–33). While tools such as HMNPPID rely on literature-based interactions and PINA 3.0 integrates RNA-seq and proteomics data, APPIC uniquely combines RNA-seq, clinical and drug repurposing data as well as survival plots. Furthermore, APPIC's real-time dynamic visualization of both 2D and 3D networks provides a more interactive and customizable experience for users compared to other tools. Unlike OncoPPi, which focuses on a curated set of cancer-specific proteins, APPIC provides a broader application across 26 cancer subtypes, making it more versatile for exploring multiple cancer types and their molecular mechanisms. A comparison of cancer PPI tools was provided in Table 1.

Limitations and future work

While APPIC offers a range of visualization options, some advanced users might find the customization options limited compared to stand-alone bioinformatics tools such as Cytoscape or Gephi, which allow for more extensive modifications and personalized network analysis workflows. To address this, all APPIC output files are available to the users, enabling them to continue their analysis and visualization on other platforms for their specific needs.

In the selection of the top *N* highly expressed genes as 'seed genes', we assume that these genes are likely part of functionally relevant to the cancer subtype. This assumption is based on the modular organization of biological networks, where highly expressed genes often cluster into disease-relevant neighborhoods of the interactome (34). While not all highly expressed genes are the most critical cancer drivers, this approach leverages network biology principles such as 'guilt by association' to prioritize genes for further analysis (35). However, we acknowledge that this is a limitation and are exploring ways to refine gene selection using multi-omics data and functional enrichment.

As a future work, APPIC will be updated with PPI network data from other cancer types such as hematological malignancies and sarcomas. Further integration of APPIC with other specialized resources in cancer biology and drug resistance could significantly enhance its impact. One particularly promising resource is the DRMref database, which provides a comprehensive map of drug resistance mechanisms based on single-cell data (36). DRMref offers insights into various cancer subtypes, including detailed analyses of cellular composition, intratumoral heterogeneity and gene expression in resistant cells. By incorporating this database, APPIC could extend its ability to examine how PPI networks are influenced by drug resistance, enabling researchers to explore the underlying mechanisms and discover novel therapeutic targets.

Conclusions

APPIC emerges as an innovative platform that not only facilitates a deeper understanding of cancer subtypes through PPI subnetworks but also acts as a catalyst for the development of precision medicine strategies. Its unique integration of multiple databases, coupled with a user-friendly interactive interface, positions APPIC as an essential tool for both researchers and clinicians working in the field of oncology.

Data availability

APPIC is freely available at https://appic.brown.edu.

Acknowledgements

We would like to thank Brown University Center for Computation and Visualization (CCV) for computational resources and Brown University Office of Information Technology for website support.

Funding

Brown University Undergraduate Teaching and Research Awards (UTRA) (Charissa C. and Caden C.); Legorreta Cancer Center, Brown University (to E.D.G.U and A.U.); Rhode Island Foundation (to E.D.G.U. and A.U.); Warren Alpert Medical School, Brown University (to B.A.).

Conflict of interest statement

None declared.

References

1. Siegel,R.L., Giaquinto,A.N. and Jemal,A. (2024) Cancer statistics, 2024. *CA Cancer J. Clin.*, **74**, 12–49.
2. Heron,M. and Anderson,R.N. (2016) Changes in the leading cause of death: recent patterns in heart disease and cancer mortality. *NCHS Data Brief*, **254**, 1–8.
3. Grechkin,M., Logsdon,B.A., Gentles,A.J. and Lee,S.I. (2016) Identifying network perturbation in cancer. *PLoS Comput. Biol.*, **12**, e1004888.
4. Liu,Z., Weng,S., Dang,Q., Xu,H., Ren,Y., Guo,C., Xing,Z., Sun,Z. and Han,X. (2022) Gene interaction perturbation network deciphers a high-resolution taxonomy in colorectal cancer. *eLife*, **11**, e81114.
5. Stratton,M.R., Campbell,P.J. and Futreal,P.A. (2009) The cancer genome. *Nature*, **458**, 719–724.
6. Gerlinger,M., Rowan,A.J., Horswell,S., Math,M., Larkin,J., Endesfelder,D., Gronroos,E., Martinez,P., Matthews,N., Stewart,A., *et al.* (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.*, **366**, 883–892.
7. Kim,S., Kim,D.H., Lee,W., Lee,Y.M., Choi,S.Y. and Han,K. (2020) The nature of triple-negative breast cancer classification and antitumoral strategies. *Genomics Inform.*, **18**, e35.
8. Oti,M., Snel,B., Huynen,M.A. and Brunner,H.G. (2006) Predicting disease genes using protein–protein interactions. *J. Med. Genet.*, **43**, 691–698.
9. Stark,C., Breitkreutz,B.J., Reguly,T., Boucher,L., Breitkreutz,A. and Tyers,M. (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.*, **34**, D535–D539.
10. Barabasi,A.L., Gulbahce,N. and Loscalzo,J. (2011) Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.*, **12**, 56–68.
11. Szklarczyk,D., Gable,A.L., Nastou,K.C., Lyon,D., Kirsch,R., Pyysalo,S., Doncheva,N.T., Legeay,M., Fang,T., Bork,P., *et al.* (2021) The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.*, **49**, D605–D612.
12. Chen,K.M., Wong,A.K., Troyanskaya,O.G. and Zhou,J. (2022) A sequence-based global map of regulatory activity for deciphering human genetics. *Nat. Genet.*, **54**, 940–949.
13. Del Toro,N., Shrivastava,A., Ragueneau,E., Meldal,B., Combe,C., Barrera,E., Perfetto,L., How,K., Ratan,P., Shirodkar,G., *et al.* (2022) The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res.*, **50**, D648–D653.
14. Herwig,R., Hardt,C., Lienhard,M. and Kamburov,A. (2016) Analyzing and interpreting genome data at the network level with ConsensusPathDB. *Nat. Protoc.*, **11**, 1889–1907.
15. Kim,C.Y., Baek,S., Cha,J., Yang,S., Kim,E., Marcotte,E.M., Hart,T. and Lee,I. (2022) HumanNet v3: an improved database of human gene networks for disease research. *Nucleic Acids Res.*, **50**, D632–D639.
16. Kotlyar,M., Pastrello,C., Ahmed,Z., Chee,J., Varyova,Z. and Jurisica,I. (2022) IID 2021: towards context-specific protein interaction analyses by increased coverage, enhanced annotation and enrichment analysis. *Nucleic Acids Res.*, **50**, D640–D647.
17. Licata,L., Briganti,L., Peluso,D., Perfetto,L., Iannuccelli,M., Galeota,E., Sacco,F., Palma,A., Nardozza,A.P., Santonico,E., *et al.* (2012) MINT, the molecular interaction database: 2012 update. *Nucleic Acids Res.*, **40**, D857–D861.
18. Mostafavi,S., Ray,D., Warde-Farley,D., Grouios,C. and Morris,Q. (2008) GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol.*, **9**(Suppl. 1), S4.
19. Peri,S., Navarro,J.D., Kristiansen,T.Z., Amanchy,R., Surendranath,V., Muthusamy,B., Gandhi,T.K., Chandrika,K.N., Deshpande,N., Suresh,S., *et al.* (2004) Human protein reference database as a discovery resource for proteomics. *Nucleic Acids Res.*, **32**, D497–D501.
20. Persson,E., Castresana-Aguirre,M., Buzzao,D., Guala,D. and Sonnhammer,E.L.L. (2021) FunCoup 5: functional association networks in all domains of life, supporting directed links and tissue-specificity. *J. Mol. Biol.*, **433**, 166835.
21. Armanious,D., Schuster,J., Tollefson,G.A., Agudelo,A., DeWan,A.T., Istrail,S., Padbury,J. and Uzun,A. (2020) Proteinarium: multi-sample protein–protein interaction analysis and visualization tool. *Genomics*, **112**, 4288–4296.
22. Hacking,S., Chou,C., Baykara,Y., Wang,Y., Uzun,A. and Gamsiz Uzun,E.D. (2023) MMR deficiency defines distinct molecular subtype of breast cancer with histone proteomic networks. *Int. J. Mol. Sci.*, **24**, 5327.
23. Cerami,E., Gao,J., Dogrusoz,U., Gross,B.E., Sumer,S.O., Aksoy,B.A., Jacobsen,A., Byrne,C.J., Heuer,M.L., Larsson,E., *et al.* (2012) The cBio Cancer Genomics Portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.*, **2**, 401–404.
24. Povey,S., Lovering,R., Bruford,E., Wright,M., Lush,M. and Wain,H. (2001) The HUGO Gene Nomenclature Committee (HGNC). *Hum. Genet.*, **109**, 678–680.
25. Raudvere,U., Kolberg,L., Kuzmin,I., Arak,T., Adler,P., Peterson,H. and Vilo,J. (2019) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.*, **47**, W191–W198.
26. Thul,P.J. and Lindskog,C. (2018) The human protein atlas: a spatial map of the human proteome. *Protein Sci.*, **27**, 233–244.
27. Feng,Y., Wang,Q. and Wang,T. (2017) Drug target protein–protein interaction networks: a systematic perspective. *Biomed. Res. Int.*, **2017**, 1289259.
28. Ghadermarzi,S., Li,X., Li,M. and Kurgan,L. (2019) Sequence-derived markers of drug targets and potentially druggable human proteins. *Front. Genet.*, **10**, 1075.
29. Sokal,R.R. and Rohlf,F.J. (1962) The comparison of dendrograms by objective methods. *Taxon*, **11**, 33–40.
30. React (2024) React, the library for web and native user interfaces. https://react.dev/, (June 2024, date last accessed).
31. Du,Y., Cai,M., Xing,X., Ji,J., Yang,E. and Wu,J. (2021) PINA 3.0: mining cancer interactome. *Nucleic Acids Res.*, **49**, D1351–D1357.
32. Li,Q., Yang,Z., Zhao,Z., Luo,L., Li,Z., Wang,L., Zhang,Y., Lin,H., Wang,J. and Zhang,Y. (2019) HMNPPID—human malignant neoplasm protein–protein interaction database. *Hum. Genomics*, **13**, 44.

33. Li,Z., Ivanov,A.A., Su,R., Gonzalez-Pecchi,V., Qi,Q., Liu,S., Webber,P., McMillan,E., Rusnak,L., Pham,C., *et al.* (2017) The OncoPPi network of cancer-focused protein–protein interactions to inform biological insights and therapeutic strategies. *Nat. Commun.*, **8**, 14356.
34. Vidal,M., Cusick,M.E. and Barabási,A.-L. (2011) Interactome networks and human disease. *Cell*, **144**, 986–998.
35. Wolfe,C.J., Kohane,I.S. and Butte,A.J. (2005) Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks. *BMC Bioinformatics*, **6**, 227.
36. Liu,X., Yi,J., Li,T., Wen,J., Huang,K., Liu,J., Wang,G., Kim,P., Song,Q. and Zhou,X. (2024) DRMref: comprehensive reference map of drug resistance mechanisms in human cancer. *Nucleic Acids Res.*, **52**, D1253–D1264.

Received: August 21, 2024. **Revised:** November 27, 2024. **Editorial Decision:** December 22, 2024. **Accepted:** January 3, 2025