BioData Mining

SOFTWARE

Open Access

CancerHubs Data Explorer: a web application for investigating mutation-enriched protein interaction hubs in human cancers

Ivan Ferrari[1,2†], Elisa Arsuffi[1,2†], Stefano Biffo[1,2,*] and Nicola Manfrini[1,2*]

†Ivan Ferrari and Elisa Arsuffi contributed equally to this work.

*Correspondence:
Stefano Biffo
biffo@ingm.org
Nicola Manfrini
manfrini@ingm.org
[1]INGM, Istituto Nazionale Genetica Molecolare Romeo ed Enrica Invernizzi, Milan, Italy
[2]Department of Biosciences, University of Milan, Milan, Italy

Abstract

Background Cancer gene discovery has traditionally relied on single-gene analyses of genomic or transcriptomic data. However, cancer is now recognized as a complex, systems-level disease driven by the coordinated dysregulation of interconnected gene sets. To harness this complexity, we recently developed *CancerHubs*, a computational framework integrating mutational profiles, clinical outcome predictions, and interactomics to prioritize genes central to cancer pathogenesis. At its core, *CancerHubs* introduces the *Network Score*, a novel metric quantifying a gene's involvement in cancer by counting the number of mutated interactors of its encoded protein, highlighting central "hub" genes likely playing pivotal roles in cancer biology.

Results Here, we present CancerHubs Data Explorer, a web-based tool for intuitive and interactive exploration of precomputed *CancerHubs* results. The application offers three main features: (i) *Gene Ranking* – to query *Network Scores* and rankings of genes across cancer types; (ii) *Subset Exploration* – to analyse curated gene subsets across selected tumours; (iii) *Network Visualization* – to explore 2D/3D interaction networks of top-ranked genes. Additionally, the application provides interactive, downloadable tables and visualizations to support hypothesis generation and gene prioritization for cancer research and precision oncology. Its intuitive and robust interface makes it suitable for both bench scientists and computational researchers, across clinical and research settings.

Conclusions The CancerHubs Data Explorer transforms a static prioritization framework into an accessible, dynamic platform that supports hypothesis generation for cancer biology and precision oncology. By integrating heterogeneous data sources into a user-friendly interface, the tool enables both computational and experimental researchers to identify functionally relevant cancer hubs without requiring coding expertise.

Keywords Cancer genomics, Gene prioritization, Protein–protein interaction networks, Hub genes, Network biology, Bioinformatics software, Precision oncology, Systems biology, Oncogene, Tumour suppressor

Background

Numerous studies have shown that cancer phenotypes—such as uncontrolled proliferation, immune evasion, and metastasis—are rarely driven by single genetic events alone. Instead, they emerge from coordinated disruptions within functional modules of the interactome [1–11], often involving central "hub" genes that act as key regulators of biological networks [12]. This systems-level complexity underscores the need for network-based approaches to fully understand cancer biology [3, 13]. Such strategies can reveal functionally important genes that may be overlooked by traditional frequency-based analyses and are therefore rapidly emerging as promising tools for identifying novel therapeutic targets.

Following this systems-level rationale, we recently developed *CancerHubs* [14], a first-of-its-kind computational workflow that integrates somatic mutation profiles, prognostic scores derived from gene expression data [15], and curated protein–protein interaction networks [16] to systematically prioritise cancer-related hub genes (Fig. 1A). *CancerHubs* prioritizes genes using the *Network Score* (Fig. 1B), a novel metric that quantifies the extent to which a gene is embedded within a cluster of mutated proteins in a given tumour type.

Despite its seminal contribution, our original implementation of *CancerHubs* had notable limitations in reaching the broader scientific community: it was static and included data for only five tumour types.

To overcome these limitations and enhance *CancerHubs* accessibility, we developed CancerHubs Data Explorer, an interactive web application that extends the original framework by incorporating additional tumour types, interactive network visualisation, and pan-cancer analyses. The platform enables users to directly explore and interpret tumour-specific mutational hubs, as well as to assess the likelihood of genes being involved across multiple or diverse cancer types through features such as the Common Genes panel and the pan-cancer score. The platform offers ranked gene tables, cross-tumour comparisons, and interactive 2D/3D network visualizations.

Designed for both usability and scientific depth, the application supports hypothesis generation by allowing users to filter genes based on mutation status and prognostic relevance, compare rankings across tumour types, and export results. By eliminating the need for coding expertise, the CancerHubs Data Explorer empowers both computational and experimental researchers to derive systems-level insights into mutation-driven network architecture and gene relevance, bridging the gap between large-scale omics data and actionable discovery in cancer research.

Implementation

The CancerHubs Data Explorer is a Shiny-based web application that makes *CancerHubs* results interactive and widely accessible. It integrates mutation profiles (Supplementary Table S1), PRECOG-derived prognostic scores, which quantify the correlation or anti-correlation between expression of a specific gene and overall survival of patients [15], and curated protein–protein interaction networks, to generate tumour-specific *Network Scores* across 11 cancer types. The interface is organized into modular panels: View Dataframe (ranked tables), Gene Ranking (cross-cancer comparisons), Common Genes (recurrence analysis), Network Plot (2D/3D interactomes), and Gene Network

(local interaction neighbourhoods). All tables and visualizations are exportable for downstream use.

CancerHubs Data Explorer is freely accessible at https://cancerhubs.app and runs on all major browsers without requiring installation or user registration. Source code and documentation is actively maintained by the INGM, Milan, Italy and can be found on GitHub (https://github.com/ingmbioinfo/cancerhubs_shiny).

Results

The CancerHubs Data Explorer enables users to filter results across four predefined gene subsets, each offering a distinct perspective on gene relevance within tumour-specific networks.

- The **All Genes** subset includes every gene scored by the *CancerHubs* framework, serving as a comprehensive network-level reference regardless of mutation status or prognostic annotation.
- The **Only Mutated** subset focuses exclusively on genes harbouring mutations in the selected tumour, highlighting candidates with direct genomic alterations.
- The **PRECOG** subset features genes with strong prognostic associations, as defined by *meta-Z* scores from the PRECOG database, regardless of their mutation status.
- The **Only PRECOG** subset isolates prognostic genes that are not mutated, potentially uncovering important regulators with functional relevance but low mutation frequency.

These categories help researchers tailor their exploration based on mutational evidence, prognostic significance, or both.

To ensure relevance, genes lacking both mutation and prognostic annotation are excluded during preprocessing and do not appear in the "All Genes" subset. This filtering step ensures that ranked outputs focus on genes with at least some supporting evidence for cancer involvement.

Each feature in the CancerHubs Data Explorer is organized into a dedicated tab within the application interface. These features are designed to support layered exploration, from broad mutational trends to focused analysis of gene-level interactions, with minimal user effort.

View Dataframe panel

The "View Dataframe" panel (Fig. 1C) is the primary interface for exploring tumour-specific rankings. Users begin by selecting a cancer type, which dynamically updates the associated table. Each gene is listed with its mutation count, PRECOG *meta-Z score*, and *Network Score*. Real-time filtering by subset allows comparative views, such as ranking mutated versus non-mutated hubs. The full table is exportable in CSV or Excel formats enabling local analyses or integration into external workflows.

Gene ranking panel

The "Gene Ranking" panel (Fig. 1D) allows users to evaluate the network importance of a specific gene across all cancer types.

Entering a gene symbol generates:

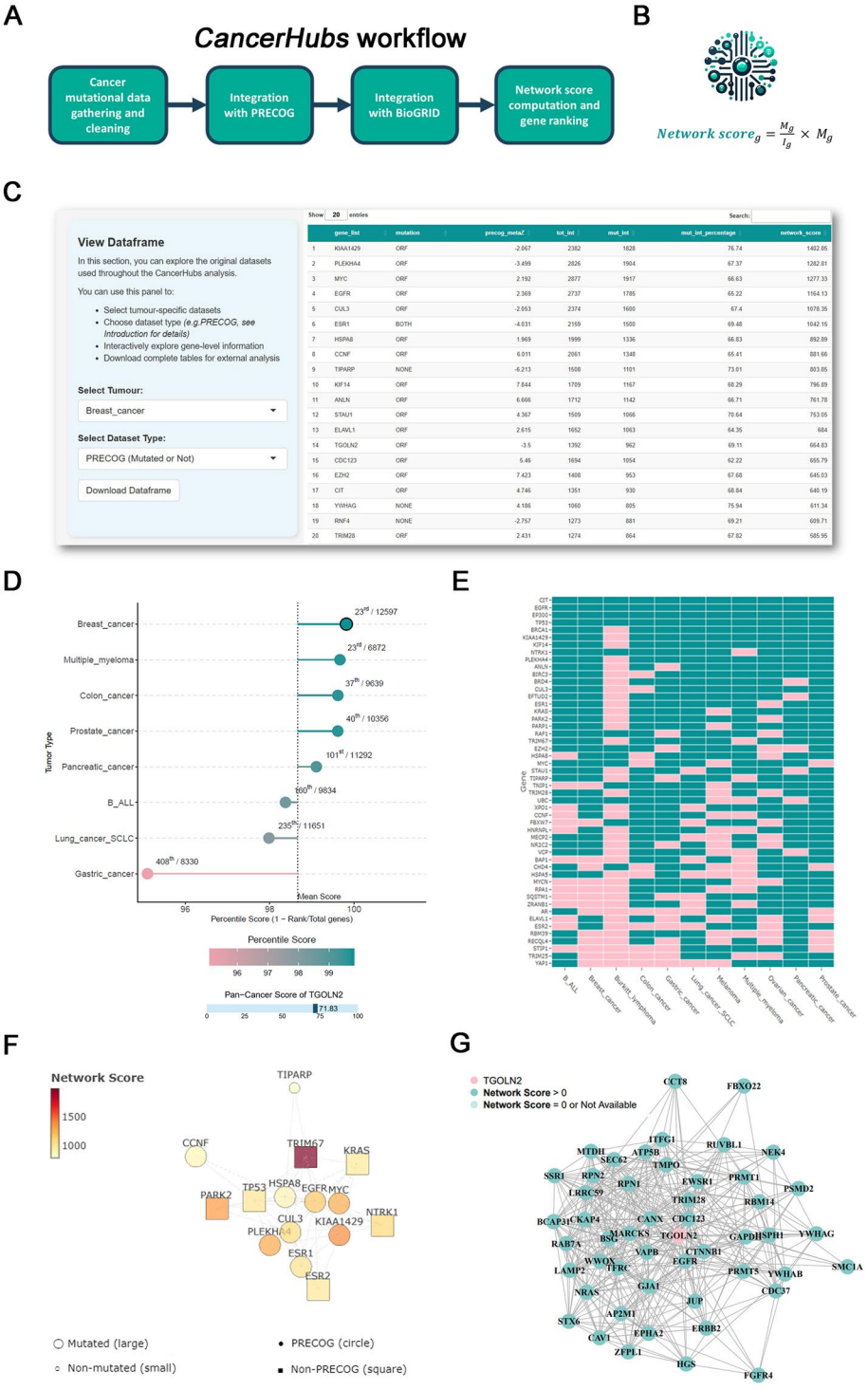

Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Examples of CancerHubs Data explorer outputs. (**A**) The *CancerHubs* workflow. The CancerHubs Data explorer is based on *CancerHubs*, a computational framework designed to identify novel cancer-related protein interaction hubs by integrating mutational data, clinical outcome predictions (*PRECOG*) and protein–protein interaction networks (*BioGRID*). The approach ultimately assigns to each gene analysed a *Network Score* which predicts its involvement in cancer. (**B**) *Network Score* formula. For a given gene "*g*" in a specific tumour, *CancerHubs* calculates the *Network Score* as the ratio between the number of mutated interactors "*M*" of the protein encoded by that specific gene and the total number of interactors "*I*", multiplied by the number of mutated interactors "*M*". (**C**) Example of an output obtained using the "View Dataframe" feature. List of the top 20 *Network scoring* genes that are also associated with clinical outcome (i.e., significant according to the *PRECOG Meta-Z* score) in breast cancer. (**D**) Example of an output obtained using the "Gene Ranking" feature. Percentile rank of the *TGOLN2* gene across different tumour types based on *Network Scores*. Higher ranks indicate greater predicted involvement in cancer. The *Pan-cancer score* represents an overall summary of the gene's rank across all analysed tumour types. (**E**) Example of an output obtained using the "Common Genes" feature. The heatmap shows those genes appearing among the top 50 ranked candidates across multiple cancers. Presence is colour-coded (pink = not present, teal = present). (**F**) Example of an output obtained using the "Network Plot" feature. The top 20 *PRECOG*-scoring genes in breast cancer are shown. (**G**) Example of an output obtained using the "Gene Network" feature. Direct interactors of TGOLN2 in breast cancer are shown. The input gene is highlighted in pink. Nodes are coloured by *Network Score* and include up to 50 interactors with optional interactor–interactor connections

1. A horizontal lollipop chart showing the gene's percentile rank per tumour, based on its *Network Score* (Fig. 1D, top). This plot provides a quick visual impression of how prominently the gene features in the network landscape across cancers.
2. A pan-cancer score summarizing the gene's overall centrality across the dataset (Fig. 1D, bottom). This score integrates information from all cancer-specific rankings to offer a more global view of the gene's involvement in cancer biology, highlighting genes that are consistently or exceptionally central in one or more cancers.

This panel also includes a table of tumour-specific scores and rankings, which can be exported along with the plot for reporting or validation.

Common genes panel

The "Common Genes" panel (Fig. 1E) identifies genes that appear repeatedly among the top-ranking candidates across tumours. Users specify how many top genes to include per cancer and how many tumour types a gene must appear in. The result is a binary heatmap showing cross-tumour recurrence, supporting prioritization of broadly relevant cancer genes.

Network plot panel

The "Network Plot" tab (Fig. 1F) provides a 3D visualization of top-ranking genes and their interactions within a selected tumour. Nodes represent genes, while edges represent known protein-protein interactions. Colouring and layout can reflect *Network Score* or PRECOG score. Users can filter to include only mutated interactors, adjust the number of top genes and export the network as edge/node tables. This visualization is useful for identifying functional modules or subclusters of genes that may participate in shared biological processes.

Gene network panel

The "Gene Network" panel (Fig. 1G) provides a focused view of a single gene's immediate interaction neighbourhood. Upon selecting a gene, the application retrieves up to 50 direct interactors from the tumour-specific dataset, ranking them by *Network Score*. Users can include only mutated interactors or simplify the network by hiding connections between interactors. This view provides researchers with insight into the

immediate molecular environment of a candidate gene, supporting hypothesis generation about its potential roles in complex formation, signalling or regulatory circuits. For users interested in larger neighbourhoods, the complete tumour-specific interactome tables are available for download directly from the same tab.

Illustrative biological use cases

To demonstrate how CancerHubs Data Explorer supports both validation of known drivers and discovery of functionally relevant network hubs, we present two worked examples: one that uses a pan-cancer approach and the other that focuses on exploring protein hubs involved in prostate cancer.

Pan-cancer use case

To explore pan-cancer protein hubs, we used the "Common Genes" panel to identify genes appearing among the top 35 *Network Score* candidates in at least 8 of the 11 cancers analysed. This was achieved by setting the "Number of top genes" option to 35 and the "Min. Presence in Tumours" option to 8 (Fig. 2A). We decided to unbiasedly analyse all genes and did so by selecting the "All genes" dataset type from the "Select Dataset Type" option (Fig. 2A). This search identified 23 recurrent hubs (Fig. 2B). As expected, this set contained classical pan-cancer drivers such as EGFR, TP53 and KRAS [17] and widely defined guardians of genome/chromatin integrity such as EP300, BRD4, PARP1 and EZH2 [18, 19], validating that the *Network Score* recovers well-established cancer-associated genes. Beyond these canonical cancer-related genes, several less canonical ones, but with putative clinical relevance, were found. These included CIT, a Serine/Threonine Kinase which regulates of cytokinesis; KIAA1429, a component of the m6A RNA methyltransferase complex recently proposed to be a pan-cancer prognosis factor [20]; PLEKHA4, a pleckstrin homology domain protein recently described as a Wnt/β-Catenin signaling modulator and promising drug target in melanoma [21]; TRIM67, an E3 ubiquitin ligase with both oncogenic [22] and tumour suppressor capabilities [23]; STAU1, an RNA-binding protein involved in mRNA transport and localization recently depicted as an oncogenesis regulator [24] and TIPARP, a mono-ADP-ribosyltransferase that acts as a key regulator in cellular signaling and that has been recently identified as a key regulator of tumour immune evasion [25]. These genes represent diverse molecular mechanisms, including RNA modification, post-translational regulation, and cell division control, highlighting the potential of network-based approaches to identify non-traditional cancer vulnerabilities that may be overlooked by frequency-based mutational analyses.

To illustrate how individual genes within these modules can be explored in detail, we focused on CIT, which ranked among the top *Network Score* hubs in all 11 tumour types, alongside EGFR, TP53, and EP300. CIT, Citron Rho-Interacting Serine/Threonine Kinase, is a cytokinesis regulator that has been associated with aggressive behaviour and poor prognosis in bladder [26], prostate [27], brain cancer [28] and multiple myeloma [29]. CIT represents a particularly informative example of the added value of the *CancerHubs* approach because it is not frequently mutated at the sequence level in cancer [13] and is absent from pan-cancer catalogues of recurrently mutated driver genes identified

Fig. 2 Pan-cancer use case. **A**) Parameter settings used in the "Common Genes" panel to identify recurrent pan-cancer hubs. The analysis was set to identify the top 35 Network Score genes present in at least 8 of the 11 tumour types, using the "All Genes" dataset. **B**) Heatmap output from the "Common Genes" panel showing the 23 recurrent genes identified. Rows represent genes and columns represent cancer types. Presence is colour-coded (pink = not present, teal = present). All listed genes are among the top 35 scorers in all the 11 tumours considered. **C**) Output from the "Gene Ranking" panel for CIT. The top bar indicates CIT's Pan-Cancer Score. The lollipop chart below shows the percentile rank of CIT in each individual tumour type. **D**) Interaction network of CIT in melanoma obtained from the "Gene Network" panel. The central node (pink) is CIT. Surrounding nodes (teal) represent its top 50 interactors, ranked by Network Score. "*CIT*" was used as gene target, "Melanoma" as cancer type and "All Genes" as dataset type

by large-scale sequencing [30], making it typically overlooked by mutation-frequency-based prioritisation strategies.

Using the "Gene Ranking" panel, we found that CIT achieved a pan-cancer score of 99.83, confirming its putative pan-cancer role and indicating exceptional network centrality across cancers, with the highest percentile score observed in melanoma (Fig. 2C).

To explore CIT's functional context, we used the "Gene Network" panel to visualize its direct interactors in melanoma using the "All Genes" Dataset type (Fig. 2D). When

specifically focusing on CIT's top 20 interactors (Supplementary Table S2), the resulting network revealed a multi-layered functional architecture organized around at least five interconnected modules:

1. A cytokinesis and cell division module, composed by ANLN and KIF14, both of which also appear among the 23 recurrent pan-cancer hubs (Fig. 2A), confirming CIT's canonical cytokinesis role.
2. A genome stability and tumour suppression module, composed by TP53 and BRCA1, demonstrating that CIT's network bridges cell division with DNA damage response and repair pathways essential for maintaining genomic integrity.
3. A proteostasis and stress adaptation module made up by recurrent pan-cancer hubs including PARK2, CUL3, HSPA8, and SQSTM1 and TRIM67, indicating tight integration with protein quality control, autophagy, and metabolic stress response.
4. An RNA processing and epitranscriptomic control module composed by KIAA1429 and EFTUD2 linking cell division to mRNA methylation and the splicing machinery.
5. A signalling and transcriptional control module composed by RAF1, PLEKHA4, EZH2 and hormone receptors ESR1 and ESR2, connecting cytokinesis to growth factor signalling, Wnt pathway modulation, epigenetic reprogramming, and hormone-responsive transcription.

Strikingly, 14 of CIT's top 20 interactors (TRIM67, PARK2, PLEKHA4, KIAA1429, CUL3, ESR1, HSPA8, TP53, ANLN, KIF14, RAF1, EFTUD2, CIT and EZH2) are themselves recurrent pan-cancer hubs identified in the Common Genes analysis (Fig. 2A). This convergence demonstrates that CIT might not function as an isolated cytokinesis regulator but rather as a central integration hub coordinating cell division with genome stability, proteostasis, RNA metabolism, and signalling networks. The consistent top ranking of CIT across all 11 tumour types, coupled with the clustering of multiple pan-cancer modules within its immediate neighbourhood, supports the hypothesis that dysregulated cytokinesis represents a network-level vulnerability that simultaneously destabilizes multiple coupled biological systems essential for cancer cell survival.

This integrated view nominates CIT as a high-priority candidate for functional validation and demonstrated how CancerHubs Data Explorer can be used to define novel, otherwise overlooked, pan-cancer genes and/or regulatory networks.

Prostate cancer use case

Concerning a specific cancer use case, a user interested in exploring prostate cancer hub genes should use the "View Dataframe" tab selecting "Prostate_cancer" from the "Tumour type" option (Fig. 3A). Specifically, to restrict the analysis those genes with a predicted impact on patients' overall survival, the user should select the "PRECOG" dataset type from the "Select Dataset type" option (Fig. 3A). The resulting output is a dataframe of genes ranked by *Network Score*, all of which are predicted to be significantly positively or negatively associated with patients' overall survival, as indicated by the "precog_metaZ" score column (Fig. 3B). When restricting the analysis to the top 25 genes by setting "25" in the "Show entries" option (Fig. 3B), the user will first encounter a reassuring validation of the *CancerHubs* framework: canonical prostate cancer drivers, namely AR (Androgen Receptor), TP53, BRCA1, and ESR2, are all present in the top 25 ranking list, confirming the tool's ability to recover the core drivers of prostate

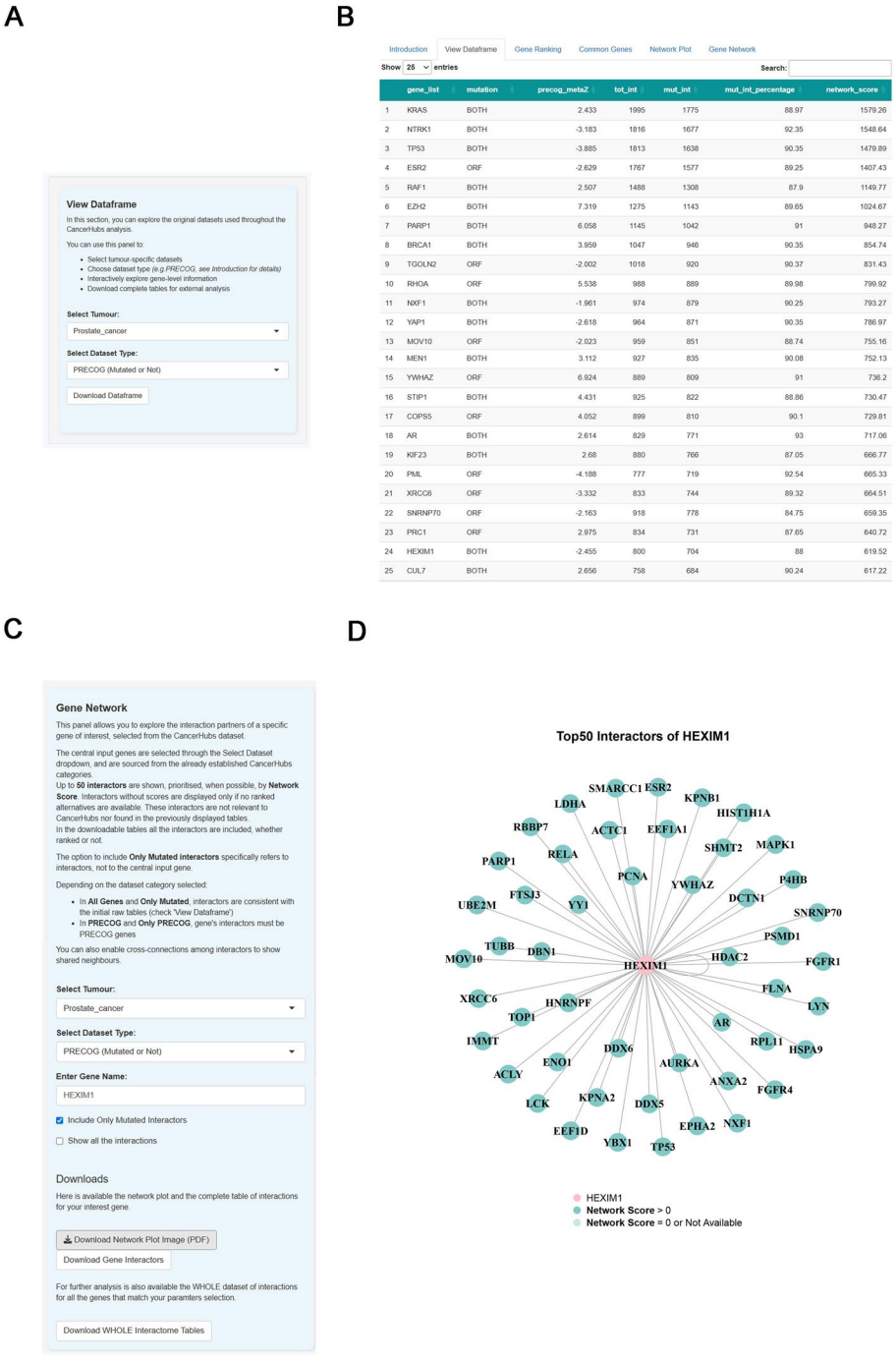

Fig. 3 Prostate cancer use case. **A**) Parameter settings used in the "View Dataframe" panel to explore prostate cancer hubs. The "Prostate_cancer" tumour type and the "PRECOG (Mutated or Not)" dataset type were selected to filter for prognostically relevant genes. **B**) Resulting dataframe showing the top 25 ranked genes in prostate cancer sorted by Network Score. Columns display: gene name, mutation location (ORF, NON_ORF, BOTH, NONE), PRECOG Meta-Z score, total number of interactors, total number of mutated interactors, percentage of mutated interactors and Network Score. **C**) Parameter settings in the "Gene Network" panel to visualize the HEXIM1 neighbourhood. We selected "HEXIM1" as the gene target to analyse and "Prostate_cancer" and the "PRECOG" dataset type to focus on prognostically significant interactors. **D**) Interaction network of HEXIM1 in prostate cancer obtained from the "Gene Network" panel. The central node (pink) is HEXIM1. Surrounding nodes (teal) represent its top 50 interactors, ranked by Network Score

tumorigenesis and hormone signalling. Besides prostate cancer-specific drivers, the user will also encounter well-established pan-cancer oncogenes, including KRAS, NTRK1, RAF1 and EZH2; tumour suppressors, including TP53, MEN1, PML and genome maintenance factors, including PARP1 and XRCC6. Their high *Network Scores* in prostate cancer suggest these ubiquitous drivers operate through interaction modules that are highly relevant to prostate cancer pathology.

The tool also surfaces genes less traditionally associated with this specific tumour such as: TGOLN2, which was however recently predicted through *CancerHubs* to be a putative pan-cancer tumour suppressor [14]; NXF1, the principal nuclear mRNA export factor, recently implicated in suppressing tumour progression by modulating alternative splicing of SP4 in endometrial cancer [31], MOV10 and SNRNP70, two RNA modulators and KIF23, a mitotic kinesin whose overexpression correlates with poor prognosis and recurrence in multiple cancer types and which is a promising chemotherapy target [32]. Besides these genes we found HEXIM1, which represents a particularly interesting hub. HEXIM1 encodes for an AR co-repressor whose expression is progressively downregulated during prostate cancer development and whose enforced re-expression has been shown to restore anti-androgen sensitivity [33].

To explore HEXIM1's interaction network in prostate cancer, we opened the "Gene Network" panel, selecting "Prostate_cancer" in the "Select Tumour" option (Fig. 3C). Again, we decided to focus only on genes with a predicted association with clinical outcome, so we selected the "PRECOG" dataset in the "Select Dataset Type" option (Fig. 3C). HEXIM1 top 50 interactors are shown in Fig. 3D.

When specifically focusing on HEXIM1's top 20 PRECOG-significant interactors (Supplementary Table S3) downloading the full interactor list using the "Download Gene Interactors" option (Fig. 3C), the resulting network was organized around four main interconnected modules:

1. A transcriptional and chromatin control module, composed by the TP53, ESR2, AR and HDAC2, demonstrating HEXIM1's direct embeddedness within the core hormone-responsive transcriptional machinery.
2. A genome maintenance and DNA repair module, composed by PARP1, XRCC6 and TOP1, connecting HEXIM1 to multiple DNA damage response and repair pathways whose integrity determines therapeutic vulnerability.
3. A post-transcriptional RNA processing and export module, composed by NXF1, SNRNP70 and MOV10, suggesting that HEXIM1 functionally couples transcriptional control with downstream RNA processing, splicing, and nuclear-cytoplasmic transport.
4. A signal transduction and survival scaffolding module, composed by YWHAZ, EPHA2 and LYN, linking HEXIM1 to cell survival, death decision nodes, and receptor-mediated signalling pathways.

This specific neighbourhood positioning suggests that HEXIM1 sits at a critical interface between AR-driven transcriptional control and multiple layers of regulation, spanning chromatin state, DNA integrity, RNA metabolism, and survival signalling.

This analysis suggests to prioritise HEXIM1 for functional validation in prostate cancer and demonstrates how CancerHubs Data Explorer can be used to define novel, otherwise overlooked, cancer specific genes and/or regulatory networks.

Discussion

CancerHubs Data Explorer extends a static, network-based prioritisation framework into a versatile and accessible resource for hypothesis generation in oncology. By integrating mutational, prognostic, and interactomic data across 11 cancer types, the application allows users to explore the tumour-specific and pan-cancer relevance of candidate genes through intuitive, multi-layered visualisations. Its filterable gene subsets, interactive ranking metrics, and network-based interfaces facilitate the identification of well-established, as well as previously underexplored, cancer-associated hubs.

Developed by computational and experimental researchers for a broad research community, the platform enables rapid and reproducible exploration of cancer-relevant networks without requiring coding expertise. Exportable outputs support seamless integration into publications, presentations, and downstream workflows.

Numerous platforms for cancer gene prioritisation currently exist, employing strategies based on mutation frequency, gene expression profiles, or pathway enrichment. Prominent examples include cBioPortal [34], IntOGen [35] and NetworkAnalyst [36]. While these tools are widely used and valuable, they typically do not integrate tumour-specific interactome topology. For instance, NetworkAnalyst offers network visualisation and computes topological measures such as centrality, but it relies on general-purpose protein–protein interaction networks that are not tailored to individual cancer types or mutational landscapes. As a result, such platforms lack the ability to quantify gene centrality within context-specific networks, an important aspect for capturing systems-level properties of tumour biology. Graphics-oriented approaches such as DriverNet [37], HotNet2 [38], and NetSig [39] do incorporate aspects of network structure but usually require raw data input and significant computational expertise, limiting their accessibility.

Other resources, such as CancerGeneNet [40] and OncoScore [41], rely on curated or text-mined associations but lack a quantitative, interactome-based scoring system.

In contrast, by relying on the *CancerHubs* approach, the CancerHubs Data Explorer combines mutational burden, prognostic relevance, and network centrality into a composite *Network Score* and delivers these results through an intuitive, Shiny-based interface.

This unique integration provides a tumour-specific, network-informed perspective that is not captured by frequency-based or generic network approaches, making it a valuable addition to the existing ecosystem of cancer genomics tools.

It is important to note that the *Network Score* is designed as a topology-centric screening metric not as a predictor of patient survival or driver status. It is computed solely from mutational and interactomic data to capture the centrality of a gene within mutation-enriched networks, independent of clinical outcome. While we integrate PRECOG scores to allow filtering by prognostic relevance, this serves as a complementary layer of information. This dual-metric design empowers users to explore genes that are either centrally embedded in mutation networks, prognostically relevant, or both, without merging network topology with survival prediction. Furthermore, by calculating an unbiased weighted count of mutated neighbours not considering interaction directionality nor regulatory logic, our approach avoids the constraints of overspecified models that often rely on incompletely characterized interactions. While this design choice prioritizes broad sensitivity by capturing genes embedded within mutation-enriched

subnetworks regardless of their specific molecular roles, it may occasionally elevate highly mutated but functionally peripheral genes. Therefore, we recommend that users view the *Network Score* as a hypothesis-generation tool and validate high-scoring candidates through orthogonal approaches, such as functional assays, clinical validation, or literature review.

Ultimately, the CancerHubs Data Explorer aims to bridge large-scale genomic data and functional discovery, accelerating the identification of key molecular players in cancer biology and supporting precision oncology.

Future developments will include expanding tumour coverage, updating mutational datasets, incorporating alternative scoring methods, and enhancing interoperability with complementary platforms to maximise analytical power. In addition, integration with clinical and drug-target databases will extend its translational relevance and strengthen its applicability.

Conclusions

The CancerHubs Data Explorer provides a robust, accessible environment for prioritizing cancer-associated hub genes through integration of mutational, prognostic, and interactomic data. By combining user-friendly design with systems-level rigor, it accelerates discovery of key molecular drivers and supports precision oncology applications.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13040-026-00539-z.

Supplementary Material 1

Acknowledgements
The authors would like to thank Simone Donizetti, from the INGM of Milan, for his help with the deployment and testing of the online version of the app.

Author contributions
IF (Investigation, Methodology, Software, Supervision, Visualization), EA (Investigation, Methodology, Software, Visualization), SB (Conceptualization, Funding acquisition, Resources, Supervision), NM (Conceptualization, Funding acquisition, Resources, Supervision).

Funding
This work was supported from AIRC under MFAG 2021 - ID. 26178 project – P.I. Manfrini Nicola, from PRIN 2022 ID 20227RT83Z to Nicola Manfrini, from Worldwide Cancer Research (WCR) grant 22–0037 to Stefano Biffo and from European Union—NextGenerationEU (PNRR M4C2-Investimento 1.4 -CN00000041-PNRR_CN3RNA_SPOKE2) to Nicola Manfrini and Stefano Biffo.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Competing interests
The authors declare no competing interests.

Availability and requirements
Project name: CancerHubs Data Explorer. Project home page: https://cancerhubs.app. Source code repository: https://github.com/ingmbioinfo/cancerhubs_shiny. Operating system(s): Platform independent, available online. Programming language: R (Shiny framework, $R \geq 4.3.0$). Other requirements: None. License: MIT License. Any restrictions to use by non-academics: No restrictions; free for academic and non-academic use.

Received: 26 August 2025 / Accepted: 4 March 2026
Published online: 30 March 2026

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell Elsevier BV. 2011;144:646–74. https://doi.org/10.1016/j.cell.2011.02.013.
2. Chuang H, Lee E, Liu Y, Lee D, Ideker T. Network-based classification of breast cancer metastasis. Mol Syst Biol [Internet]. 2007 [cited 2025 Jul 9];3. https://doi.org/10.1038/msb4100180.
3. Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet Springer Sci Bus Media LLC. 2011;12:56–68. https://doi.org/10.1038/nrg2918.
4. Zhang T, Zhang D. Integrating omics data and protein interaction networks to prioritize driver genes in cancer. Oncotarget Impact Journals. 2017;8:58050–60. https://doi.org/10.18632/oncotarget.19481.
5. Shi Z, Lei JT, Elizarraras JM, Zhang B. Mapping the functional network of human cancer through machine learning and pan-cancer proteogenomics. Nat Cancer Springer Sci Bus Media LLC. 2024;6:205–22. https://doi.org/10.1038/s43018-024-00869-z.
6. Zhu Y, Zhang H, Yang Y, Zhang C, Ou-Yang L, Bai L, et al. Discovery of pan-cancer related genes via integrative network analysis. Briefings in Functional Genomics. Oxford University Press (OUP) 2022;21:325-38. https://doi.org/10.1093/bfgp/elac012.
7. Qiu J, Chen K, Zhong C, Zhu S, Ma X. Network-based protein–protein interaction prediction method maps perturbations of cancer interactome. PLoS Genet. 2021;17:e1009869. https://doi.org/10.1371/journal.pgen.1009869.
8. Li Y, Burgman B, Khatri IS, Pentaparthi SR, Su Z, McGrail DJ, et al. e-MutPath: computational modeling reveals the functional landscape of genetic mutations rewiring interactome networks. Nucleic Acids Research. Oxford University Press (OUP); 2021;49:e2–2. https://doi.org/10.1093/nar/gkaa1015.
9. Grechkin M, Logsdon BA, Gentles AJ, Lee SI. Identifying network perturbation in cancer. PLoS Comput Biol. 2016;12:e1004888. https://doi.org/10.1371/journal.pcbi.1004888.
10. Du Y, Cai M, Xing X, Ji J, Yang E, Wu J. PINA 3.0: mining cancer interactome. Nucleic Acids Research. Oxford University Press (OUP); 2021;49:D1351–7. https://doi.org/10.1093/nar/gkaa1075.
11. Sharifi Tabar M, Francis H, Yeo D, Bailey CG, Rasko JEJ. Mapping oncogenic protein interactions for precision medicine. Intl J Cancer Wiley. 2022;151:7–19. https://doi.org/10.1002/ijc.33954.
12. Jonsson PF, Bates PA. Global topological features of cancer proteins in the human interactome. Bioinformatics. Oxford University Press (OUP) 2006;22:2291–7. https://doi.org/10.1093/bioinformatics/btl390.
13. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. Sci Am Association Advancement Sci (AAAS). 2013;339:1546–58. https://doi.org/10.1126/science.1235122.
14. Ferrari I, De Grossi F, Lai G, Oliveto S, Deroma G, Biffo S et al. CancerHubs: a systematic data mining and elaboration approach for identifying novel cancer-related protein interaction hubs. Briefings in Bioinformatics [Internet]. Oxford University Press (OUP); 2024 [cited 2025 Jul 9];26. https://doi.org/10.1093/bib/bbae635.
15. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med Springer Sci Bus Media LLC. 2015;21:938–45. https://doi.org/10.1038/nm.3909.
16. Oughtred R, Rust J, Chang C, Breitkreutz B, Stark C, Willems A, et al. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Sci Wiley. 2021;30:187–200. https://doi.org/10.1002/pro.3978.
17. Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, et al. A compendium of mutational cancer driver genes. Nat Rev Cancer. 2020;20:555–72. https://doi.org/10.1038/s41568-020-0290-x.
18. Li Q, Qian W, Zhang Y, Hu L, Chen S, Xia Y. A new wave of innovations within the DNA damage response. Sig Transduct Target Ther. 2023;8:338. https://doi.org/10.1038/s41392-023-01548-8.
19. Chen Z, Tyler JK. The Chromatin Landscape Channels DNA Double-Strand Breaks to Distinct Repair Pathways. Front Cell Dev Biol. 2022;10:909696. https://doi.org/10.3389/fcell.2022.909696.
20. Ma C, Zheng Q, Wang Y, Li G, Zhao M, Sun Z. Pan-cancer analysis and experimental validation revealed the m6A methyltransferase KIAA1429 as a potential biomarker for diagnosis, prognosis, and immunotherapy. Aging. 2023;15:8664–91. https://doi.org/10.18632/aging.204968.
21. Shami Shah A, Cao X, White AC, Baskin JM. PLEKHA4 Promotes Wnt/β-Catenin Signaling–Mediated G1–S Transition and Proliferation in Melanoma. Cancer Res. 2021;81:2029–43. https://doi.org/10.1158/0008-5472.CAN-20-2584.
22. Jiang J, Ren H, Xu Y, Wudu M, Wang Q, Liu Z, et al. TRIM67 Promotes the Proliferation, Migration, and Invasion of Non-Small-Cell Lung Cancer by Positively Regulating the Notch Pathway. J Cancer. 2020;11:1240–9. https://doi.org/10.7150/jca.38286.
23. Wang S, Zhang Y, Huang J, Wong CC, Zhai J, Li C, et al. TRIM67 Activates p53 to Suppress Colorectal Cancer Initiation and Progression. Cancer Res. 2019;79:4086–98. https://doi.org/10.1158/0008-5472.CAN-18-3614.
24. Almasi S, Jasmin BJ. The multifunctional RNA-binding protein Staufen1: an emerging regulator of oncogenesis through its various roles in key cellular events. Cell Mol Life Sci. 2021;78:7145–60. https://doi.org/10.1007/s00018-021-03965-w.
25. Gozgit JM, Vasbinder MM, Abo RP, Kunii K, Kuplast-Barr KG, Gui B, et al. PARP7 negatively regulates the type I interferon response in cancer cells and its inhibition triggers antitumor immunity. Cancer Cell. 2021;39:1214–e122610. https://doi.org/10.1016/j.ccell.2021.06.018.
26. Deng P, Zhong C, He P, Liu Q, Peng S, Liu R, et al. Uncovering citron kinase as a key biomarker for predicting outcomes and therapy efficacy in non-muscle-invasive bladder cancer. Gene. 2025;963:149590. https://doi.org/10.1016/j.gene.2025.149590.
27. Rawat C, Ben-Salem S, Singh N, Chauhan G, Rabljenovic A, Vaghela V, et al. Prostate Cancer Progression Relies on the Mitotic Kinase Citron Kinase. Cancer Res. 2023;83:4142–60. https://doi.org/10.1158/0008-5472.CAN-23-0883.
28. Pallavicini G, Sgrò F, Garello F, Falcone M, Bitonto V, Berto GE, et al. Inactivation of Citron Kinase Inhibits Medulloblastoma Progression by Inducing Apoptosis and Cell Senescence. Cancer Res. 2018;78:4599–612. https://doi.org/10.1158/0008-5472.CAN-17-4060.
29. Sahin I, Kawano Y, Sklavenitis-Pistofidis R, Moschetta M, Mishima Y, Manier S, et al. Citron Rho-interacting kinase silencing causes cytokinesis failure and reduces tumor growth in multiple myeloma. Blood Adv. 2019;3:995–1002. https://doi.org/10.1182/bloodadvances.2018028456.

30. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018;173:371–e38518. https://doi.org/10.1016/j.cell.2018.02.060.
31. Meng N, Li L, Yang X, Tang H, Wang J, Zhang S, et al. NXF1 suppresses progression of endometrial cancer by interacting with the SRSF3 to regulate SP4 splicing. iScience. 2025;28:113113. https://doi.org/10.1016/j.isci.2025.113113.
32. Owida HA, Mohammad SI, Vasudevan A, Bishoyi AK, RenukaJyothi S, Panigrahi R, et al. Kinesin superfamily proteins in cancer: unveiling their role in chemotherapy. Int Immunopharmacol. 2025;166:115621. https://doi.org/10.1016/j.intimp.2025.115621.
33. Yeh I-J, Song K, Wittmann BM, Bai X, Danielpour D, Montano MM. HEXIM1 plays a critical role in the inhibition of the androgen receptor by anti-androgens. Biochem J. 2014;462:315–27. https://doi.org/10.1042/BJ20140174.
34. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal [Internet]. American Association for the Advancement of Science (AAAS); 2013. [cited 2025 Jul 9];6. https://doi.org/10.1126/scisignal.2004088.
35. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, et al. IntOGen-mutations identifies cancer drivers across tumor types. Nat Methods Springer Sci Bus Media LLC. 2013;10:1081–2. https://doi.org/10.1038/nmeth.2642.
36. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Research. Oxford University Press (OUP); 2019;47:W234–41. https://doi.org/10.1093/nar/gkz240.
37. Bashashati A, Haffari G, Ding J, Ha G, Lui K, Rosner J, et al. DriverNet: uncovering the impact of somatic driver mutations on transcriptional networks in cancer. Genome Biol [Internet] Springer Sci Bus Media LLC. 2012. https://doi.org/10.1186/gb-2012-13-12-r124. [cited 2025 Jul 9];13.
38. Reyna MA, Leiserson MDM, Raphael BJ. Hierarchical HotNet: identifying hierarchies of altered subnetworks. Bioinformatics. Oxford University Press (OUP); 2018;34:i972–80. https://doi.org/10.1093/bioinformatics/bty613.
39. Horn H, Lawrence MS, Chouinard CR, Shrestha Y, Hu JX, Worstell E, et al. NetSig: network-based discovery from cancer genomes. Nat Methods Springer Sci Bus Media LLC. 2018;15:61–6. https://doi.org/10.1038/nmeth.4514.
40. Iannuccelli M, Micarelli E, Surdo PL, Palma A, Perfetto L, Rozzo I, et al. CancerGeneNet: linking driver genes to cancer hallmarks. Nucleic Acids Research. Oxford University Press (OUP); 2020;48:D416–21. https://doi.org/10.1093/nar/gkz871.
41. Piazza R, Ramazzotti D, Spinelli R, Pirola A, De Sano L, Ferrari P, et al. OncoScore: a novel, Internet-based tool to assess the oncogenic potential of genes. Sci Rep [Internet] Springer Sci Bus Media LLC. 2017. https://doi.org/10.1038/srep46290. [cited 2025 Jul 9];7.

Publisher's note