

## **An evidence-based network approach to recommending targeted cancer therapies**

Jayaram Kancherla\*, Shruti Rao\*, Krithika Bhuvaneshwar, Rebecca B. Riggins, Robert A. Beckman, Subha Madhavan, Héctor Corrada Bravo, Simina M. Boca

April 10, 2018

\* These authors contributed equally

### **Abstract**

In this work, we introduce CDGnet, an evidence-based network approach for recommending targeted cancer therapies, available as a user-friendly informatics tool. Our approach can be used to expand the range of options of targeted therapies for cancer patients who undergo molecular profiling. It considers biological pathway information specifically by looking at downstream targets of oncogenes and is personalized for individual patients via the user-inputted molecular alterations and cancer type. CDGnet integrates disparate sources of knowledge and provides results in a number of easily-accessible and usable forms, while separating targeted cancer therapies into categories in an evidence-based manner.

### **Introduction**

In today's era of cancer precision medicine, therapeutic interventions are often tailored to an individual's tumor molecular profile, in addition to traditional considerations including age, sex, cancer stage, medical and treatment history. The term "molecular profiling" is often used to refer to some test that considers one or more biomarkers. These biomarkers may be either genetic characteristics or mRNA or protein expression values. Genetic characteristics include point mutations, insertions, deletions, duplications, gene fusions and rearrangements. They may be either germline (inherited, present in normal tissue) or somatic (present in cancer cells but not normal tissue). Expression values refer to the expression of mRNA or protein in tumors, either in comparison to other tumors or to adjacent normal tissue. Typically, tumor molecular profiling is used when a patient has few or no standard treatment options left. However, for some tumor types, it is now routine to check for specific molecular features to decide on a targeted treatment plan. For example, KRAS-wild type colorectal cancer is generally treated with EGFR inhibitors<sup>1</sup>, ER-positive (ER+) breast cancer with aromatase inhibitors or antiestrogens such as tamoxifen or fulvestrant, and HER2-positive breast cancer with monoclonal antibodies trastuzumab and pertuzumab, tyrosine kinase inhibitors such as neratinib, or antibody-toxin conjugates such as trastuzumab-DM1.<sup>2</sup> In many cases, if there is no FDA approved targeted therapy for a specific tumor type, clinicians may either recommend an off-label therapy that is prescribed for their alteration in another tumor type, or enrollment in precision medicine clinical trials (e.g. basket, umbrella, targeted therapy trials).

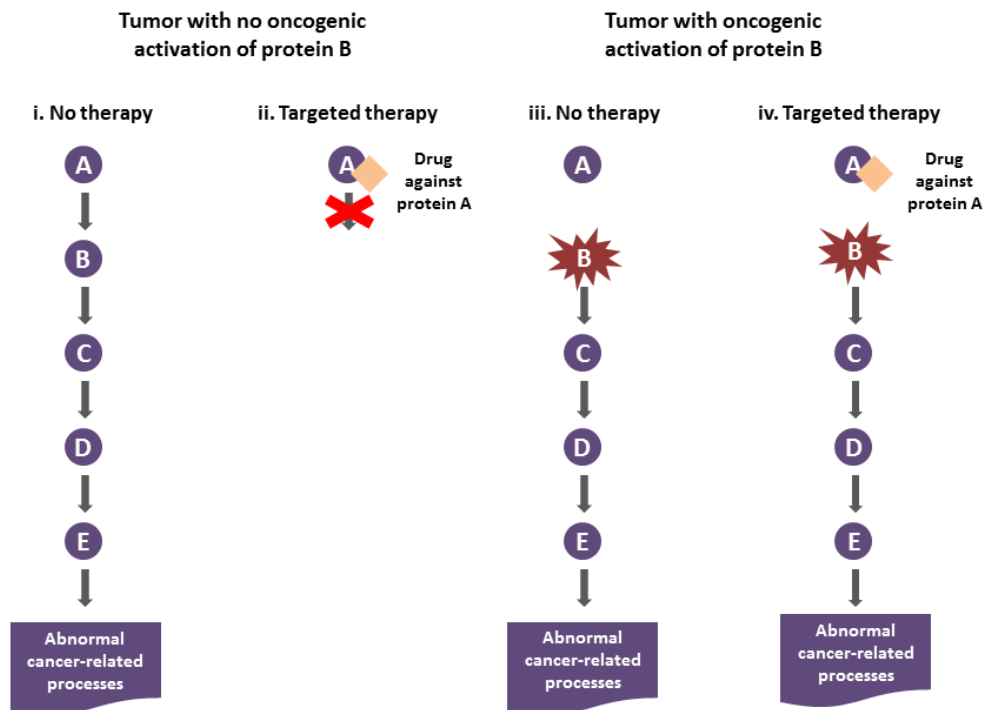
To make such decisions about off-label therapy recommendations, clinicians have to sift through vast amounts of literature and clinical databases to determine the clinical utility of variants identified through molecular profiling to decide on the appropriate treatment option for their patients. The same is true for clinical translational scientists considering relevant therapeutic approaches to evaluate in model systems or in humans, either utilizing single agents or combinations. In this setting, the number of possible molecular profiles that may be relevant and the number of experimental agents creates a combinatorial explosion of research possibilities among which prioritization is needed. Several efforts are ongoing to capture, standardize and share clinically relevant variants identified through molecular diagnostic tests between several public, academic, and private institutions,<sup>3-5</sup> though challenges remain with synthesizing evidence in a manner that is both systematic and timely<sup>6,7</sup>. Our goal in this work is to expand the range of options of targeted therapies for cancer patients who undergo molecular profiling by developing CDGnet (Cancer-Drug-Gene network), a user-friendly, evidence-based approach that accounts for downstream effects within pathways in cancer and is personalized for the individual patient. Our tool, which uses the shiny framework with an R backend<sup>8</sup>, is available at: <http://epiviz.cbcb.umd.edu/shiny/CDGnet/>. We incorporate pathway information specifically by looking at downstream targets of oncogenes, which are genes that are constitutively activated in cancer<sup>9</sup>. This is explained in the cartoon in Figure 1: If an oncogene in a biological pathway is activated, targeting genes and proteins that are found upstream may no longer be effective, leading to a focus on downstream targets. This includes the scenario of EGFR inhibitors for KRAS-wild type colorectal tumors: The EGFR protein triggers a signaling cascade in cancer, which may be blocked by anti-EGFR drugs; however, this is only effective if KRAS, which is downstream of EGFR, is not mutated. This leads to a lack of effectiveness of therapies that block EGFR. As a result, colorectal cancer patients are typically tested for KRAS mutations and EGFR inhibitors are only prescribed to individuals without specific KRAS mutations in codons 12 and 13. A comprehensive characterization of untreated colorectal tumors estimated that 43% of non-hypermutated tumors had KRAS mutations and these mutations were generally oncogenic activating mutations<sup>10</sup>, which means that a large percent of colorectal cancer patients are left with few therapeutic options. Our framework and tool are seeking to remedy this issue.

## Methods

### Overview of methods for generating patient-specific networks

The user inputs into CDGnet are the specific alterations found in a patient's tumor and the patient's cancer type. Part of the landing page is shown in **Figure 2**. These data are then integrated with: biological networks relevant to the cancer type (from the KEGG database<sup>11</sup>), FDA-approved targeted cancer therapies and indications (curated from DailyMed therapy labels<sup>12</sup>), additional gene-drug connections in the form of drug targets (from the DrugBank database<sup>13</sup>), information on whether a gene is an oncogene i.e. overactivated in cancer (from KEGG). Users may consider different data sources by using the code at <https://github.com/SiminaB/CDGnet/> directly, for example by considering the oncogenes from a recent

**Figure 1:** Simplified cartoon showing the reasoning behind looking at downstream targets of activated oncogenes. A simple pathway is shown that consists of 5 proteins, A, B, C, D, E, with A activating B, B activating C, etc, with the final activation of E leading to various abnormal cancer-related processes. Panels i and ii show the scenario where a tumor has no oncogenic activation of protein B, whereas panels iii and iv show the scenario where protein B has gained an oncogenic mutation that renders it constitutively active. If there is no oncogenic activation of protein B, then targeting protein A, as in panel ii, may be effective in stopping cancer growth. However, if there is oncogenic activation of protein B, this means that, in particular, it is not necessary for protein A to activate protein B, so that targeting protein A is not effective for turning off the pathway.



comprehensive characterization of The Cancer Genome Atlas (TCGA) projects<sup>14</sup>. Currently, the biological networks we consider are the cancer-specific pathways in KEGG and thus, for now we are also restricting the cancer types to those that have KEGG pathways. We have developed 4 different therapy categories that can be prioritized for a patient, given their specific tumor alterations, ordered from “most evidence that therapy works” to “least evidence that therapy works”:

1. FDA-approved drugs for which the input genes/proteins are biomarkers for their tumor type;
2. FDA-approved drugs for which the input genes/proteins are biomarkers in other tumor types;
3. Drugs which have as targets the input genes/proteins or as biomarkers/targets other genes/proteins that are downstream of the input oncogenes when considering the pathway corresponding to this tumor type; and
4. Drugs which have as biomarkers/targets other genes/proteins that are downstream of input oncogenes when considering the pathways corresponding to other tumor types.

In categories 3 and 4, users have the option to consider only FDA-approved targeted cancer therapies, all

**Figure 2:** Part of the landing page, which shows how users can select the cancer type and either input a tab-separated or comma-separated file or use the example data. The inset shows how under “Filter Recommended Therapies,” combinations of the first 2 checkboxes lead to the 4 different categories of therapy recommendations described below. Removing one or both of the last 2 checkboxes expands the range of therapies in Categories 3 and 4 beyond FDA-approved drugs, respectively FDA-approved targeted cancer drugs.

#### Therapy recommendations using biological networks

**Warning!** The following tool is for research purposes only. It is not intended for clinical care.

The screenshot shows a web application interface. On the left, there is a form with a dropdown menu for "Select cancer type" set to "Colorectal cancer". Below it is a file upload section with a "Browse..." button and a "No file selected" message. A table lists example gene alterations: KRAS (mutation, G13V), PIK3CA (mutation, G1049S), and BRCA2 (mutation, deleterious). Below the table is a "Filter Recommended Therapies" section with four checkboxes: "Same Cancer Type" (checked), "Same Alteration" (checked), "FDA Approved Drugs" (checked), and "FDA Approved Targeted Cancer Drugs" (checked). An arrow points from the first two checked boxes to the text "Category 1 therapies".

On the right, there is a search bar and a "Note" section. The note says "Category 1: FDA-approved drugs for which alterations in these genes/proteins are approved biomarkers in this tumor type: Show 10 entries". Below the note is a table with one entry: "1 There are no recommended therapies in this category." At the bottom right of the note section are "Previous", "1", and "Next" buttons.

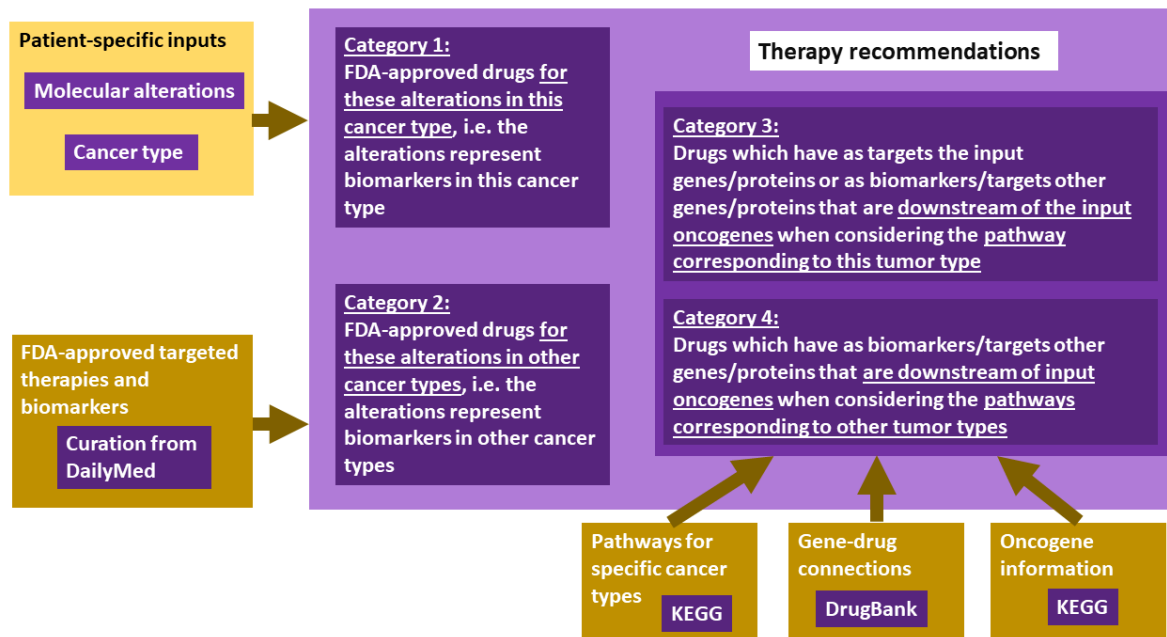
In the center, there is a blue-bordered box containing four filter combinations, each with an arrow pointing to a category name:

- Same Cancer Type → Category 2 therapies
- Same Alteration → Category 2 therapies
- Same Cancer Type → Category 3 therapies
- Same Alteration → Category 3 therapies
- Same Cancer Type → Category 4 therapies
- Same Alteration → Category 4 therapies

FDA-approved therapies, or all drugs in DrugBank; this allows for clinical researchers to consider increasing numbers of therapies only as needed, as opposed to being overwhelmed with a huge number of therapies from the start. For the clinician, this list reflects not only level of evidence but also practicality of obtaining the drug for use in their patient. Category 1 drugs will be readily available. Category 2 drugs will be readily available but may not be covered by the patient’s insurance, or may require considerable effort or justification to obtain coverage. Drugs that are not FDA approved may not be available unless their manufacturer has a compassionate use program, and then only with considerable effort.

We differentiate between targets and biomarkers because in many cases, due to complicated biological interactions, a therapy’s target may be different from the biomarker used to specify the indication, such as the case specified above with EGFR inhibitors being given for KRAS wild-type colorectal tumors, or CDK4/6 inhibitors being given for ER+ breast tumors. The general approach is presented in **Figure 3**. The options used on the landing page to obtain the different therapy categories are shown in **Figure 2**.

**Figure 3:** General approach for targeted therapy recommendations, including specific data sources.

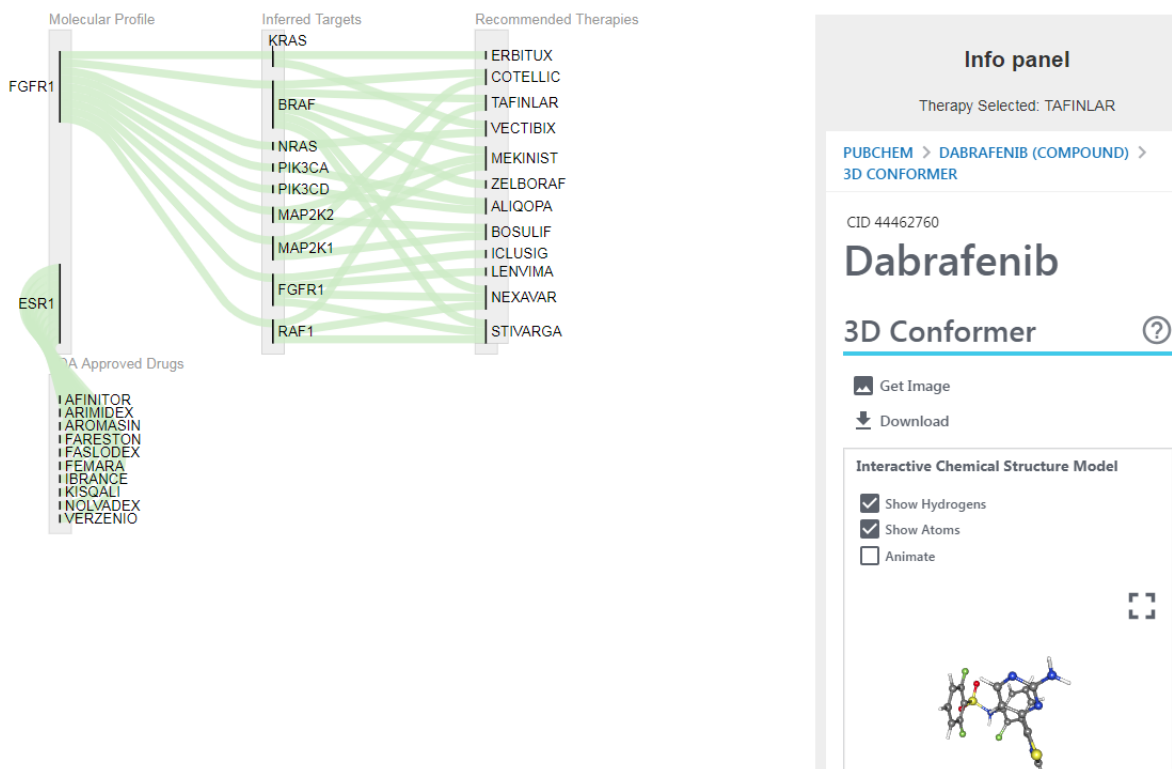


Supplementary Table 1 gives the list of FDA-approved targeted cancer therapies and indications which was obtained by considering the targeted therapies listed by the NCI<sup>15</sup> and looking up the corresponding labels at DailyMed<sup>12</sup>. In particular, the “Indications and usage” portion of the label was used to obtain the specific cancer type and the biomarker information, which is listed in the “Gene/Protein,” “Data type,” and “Alteration” columns; in the case of multiple biomarkers, these are listed on separate rows of the table. In cases where the biomarker indication is unclear, the lists of FDA companion diagnostic tests were also consulted<sup>16,17</sup>. Note that while some targeted therapies have specific biomarker indications, many do not. For example, ibrutinib (Imbruvica) is a targeted therapy, given for a number of subtypes of leukemia/lymphoma, but not for a specific indication. If there is no biomarker indication, this is noted as a “\*” in the table under the “Gene/Protein” column. The therapies are then cross-referenced with DrugBank to obtain the targets for both the therapies with biomarker indications and those without indications. The biomarkers and targets obtained in these ways are checked against downstream targets from KEGG cancer-specific pathways – which were downloaded, parsed, and with identifiers converted using the KEGGREST,<sup>18</sup> KEGGgraph,<sup>19</sup> and org.Hs.eg.db<sup>20</sup> Bioconductor packages respectively – and the information input by the user, with the gene/protein names being normalized via the rDGldb package, which is a wrapper for the DGI database<sup>21,22</sup>.

In order to obtain the list of FDA-approved drugs, we used the data files from the official Drugs@FDA resource.<sup>23</sup> Drugs@FDA contains several tab separated files (tsv) that include information on the

submission, review and approval process for various drugs. We use the `products` (list of all drugs) and `submission` (review process for all drugs) files from these files to filter for only drugs that are “Approved”

**Figure 4:** Sankey flow diagram, focusing on the flow of evidence between drug-gene and gene connections for a putative patient with ER+ breast cancer and FGFR1 overexpression, showing Category 3 recommendations, namely targets downstream of FGFR1. Therapies can be clicked to obtain a panel with PubChem information.



(AP) or “Tentatively Approved” (TA). The Drugs@FDA resource contains a list of all drugs approved since 1939 and some of these might be discontinued. As a result, we use the `marketingstatus` file to remove any discontinued products from the list. The R scripts to parse and filter the Drugs@FDA data files are available in our GitHub repository.

### Shiny app and visualization

For each of the 4 categories detailed above, a sortable and searchable table of therapies is output with the FDA-approved indications; for categories 3 and 4 network visualizations are also shown. **Figure 4** shows a Sankey flow diagram representation which focuses on the flow of evidence between drug-gene and gene-gene connections, enabling an intuitive visualization from the molecular profile to the inferred targets and recommended therapies. **Figure 5** shows a portion of the sortable and searchable corresponding table. The path column represents the pathway between the altered gene/protein and the gene/protein that is a biomarker or target; the alteration column represents the biomarker for an FDA-approved indication, if this exists, in which case the tumor for which it is approved is also listed; the

predicted effect column is “sensitive” if the alteration column is not empty, and “target” if the drug targets the protein according to the DrugBank data.

**Figure 5:** Part of the sortable, searchable table for therapies in Category 3 for a putative patient with ER+ breast cancer and FGFR1 overexpression, showing the subset of therapies that target MAP2K1.

Show 10 entries						Search: map2k1	
Drug	Gene or Protein	Type	Alteration	Path	Tumor in which it is approved	Predicted effect	
29	BOSULIF	MAP2K1	mutation	fusion (in ABL1)	FGFR1 -> SHC2 -> GRB2 -> SOS1 -> HRAS -> ARAF -> MAP2K1	Chronic myeloid leukemia	sensitive
30	COTELLIC	MAP2K1	mutation	V600K (in BRAF)	FGFR1 -> SHC2 -> GRB2 -> SOS1 -> HRAS -> ARAF -> MAP2K1	Melanoma	sensitive
31	COTELLIC	MAP2K1	mutation	V600E (in BRAF)	FGFR1 -> SHC2 -> GRB2 -> SOS1 -> HRAS -> ARAF -> MAP2K1	Melanoma	sensitive
42	MEKINIST	MAP2K1	mutation	V600E (in BRAF)	FGFR1 -> SHC2 -> GRB2 -> SOS1 -> HRAS -> ARAF -> MAP2K1	Non-small cell lung cancer	sensitive
44	MEKINIST	MAP2K1	mutation	V600E (in BRAF)	FGFR1 -> SHC2 -> GRB2 -> SOS1 -> HRAS -> ARAF -> MAP2K1	Melanoma	sensitive
46	MEKINIST	MAP2K1	mutation	V600K (in BRAF)	FGFR1 -> SHC2 -> GRB2 -> SOS1 -> HRAS -> ARAF -> MAP2K1	Melanoma	sensitive
48	MEKINIST	MAP2K1	mutation	V600E (in BRAF)	FGFR1 -> SHC2 -> GRB2 -> SOS1 -> HRAS -> ARAF -> MAP2K1	Anaplastic thyroid cancer	sensitive

Showing 1 to 7 of 7 entries (filtered from 63 total entries) Previous  Next

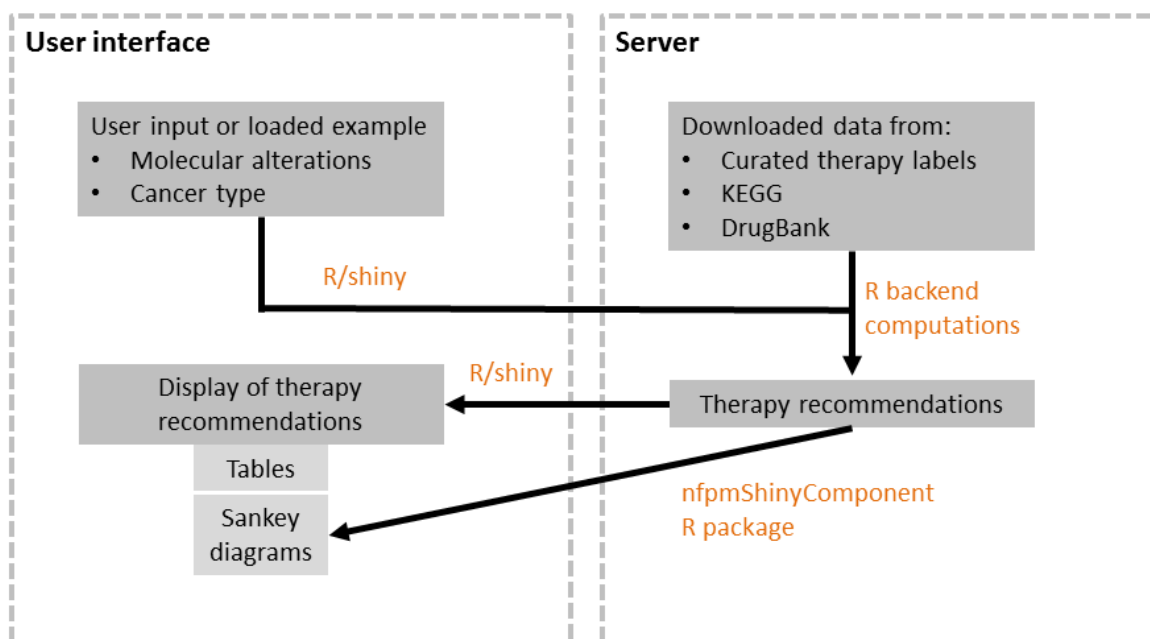
An architecture diagram for our system is shown in **Figure 6**. We use shiny, an R package/framework for creating interactive and standalone web applications directly from R<sup>8</sup>. Shiny applications can run on a webpage or can be embedded in RMarkdown documents to build interactive dashboards. They use the same technology that powers web applications - HTML and JavaScript - and allow users to create intuitive and interactive user interfaces and prototypes with an R computational backend.

To support interactive Sankey charts inside shiny applications, we developed a shiny web component for visualizing Sankey flow diagrams, available to download as an R package at <https://github.com/jkanche/nfpmShinyComponent>. Web components are custom HTML elements that are natively extensible, reusable and can be integrated with any framework that supports HTML. The Sankey visualization uses a custom three column layout to organize nodes in the graph – 1) Molecular Profile and FDA approved drugs 2) Inferred Targets and 3) Recommended Therapies – and intuitively focuses the user on the flow of evidence from input parameters to recommended therapies. The Sankey visualization also contains an information panel that displays evidence related to a pathway connection or a drug when a user selects/clicks on an edge or node. Selecting an edge shows the downstream pathway information used for inference. Selecting a recommended therapy displays the structure of the drug and linked publications from PubChem,<sup>24</sup> using PubChem widgets. The Sankey visualization is built on top of d3.js<sup>25</sup>, a data visualization library for JavaScript to build highly customizable and interactive visualizations.

## Results

We will now consider the scenario of a patient who has ER+ breast cancer. ER+ breast cancer, generally treated with aromatase inhibitors or antiestrogens, employs an array of mechanisms that permit escape from these therapies. These include amplification or upregulation of fibroblast growth factor receptor 1

**Figure 6.** Architecture diagram for our system.



(FGFR1), which is amplified in ~13% of ER+ tumors from The Cancer Genome Atlas<sup>26-28</sup> and leads to ligand-independent ER activation.<sup>29</sup> FGFR activity has also recently been shown to confer resistance to CDK4/6 inhibitors in ER+ breast cancer.<sup>30</sup> Pan-FGFR antagonists have been combined with endocrine therapies in prior clinical studies (e.g. CTKI258A2210), but the efficacy of this combination was minimal, even in patients pre-selected for alterations in the FGFR pathway.<sup>31</sup> A potential underlying explanation for this lack of benefit is that FGFR alterations impinge upon downstream signaling networks shared by many other receptor tyrosine kinases. **Figure 4** shows CDGnet recommendations for a breast cancer patient with overexpression of both ESR1 (gene encoding ER) and FGFR1, when considering only FDA-approved targeted therapies. Therapy recommendations include PIK3CA, MAPK, and RAF inhibitors, which may have utility in this context, along with the standard targeted therapies prescribed for ER+ breast cancer. **Figure 5** shows the subset of the corresponding table that consists of FDA-approved MAP2K1 inhibitors, which are approved for either ABL1 fusions or specific BRAF mutations in chronic myeloid leukemia, respectively melanoma, non-small cell lung cancer, and anaplastic thyroid cancer.



## Discussion

We developed the CDGnet tool, using an approach that considers biological pathways and connections between genes, proteins, and drugs, to prioritize targeted therapies for cancer patients. Our approach integrates many disparate sources of knowledge and provides results in an easily-accessible and usable format. Using our tool, clinicians and clinical researchers are able to quickly obtain information on the FDA approved therapies (Category 1) and potential off-label therapies (Category 2) associated with a patient's molecular profile. Our definition of Categories 1 and 2 in CDGnet are in alignment with the Tier I and II evidence level classifications recommended by the Association for Molecular Pathology (AMP), American College of Medical Genetics and Genomics (ACMG), American Society of Clinical Oncology (ASCO), and College of American Pathologists<sup>32</sup>. However, CDGnet's Categories 3 and 4 are unique to our evidence-based network approach and enable clinicians and clinical researchers to evaluate additional targeted therapy options based on their cancer patient's profile. It is important to note that the targeted therapy recommendations in Categories 3 and 4 have lower evidence levels and may or may not have proven clinical significance in ongoing clinical trials. However, by examining the downstream targets of candidate biomarkers, clinical researchers can derive key insights on potential biological pathways that can be targeted by different cancer therapies. Based on the level of evidence, the clinical actionability of these pathways can be further tested in a laboratory or clinical trial setting. Additionally, there is a growing field of research related to drug-target interactions and drug repositioning using network-based models,<sup>33–36</sup> which may in the future be integrated with our tool.

We aim to further enhance the data that drives the CDGnet tool by incorporating relevant information from additional precision oncology efforts, tools and resources. Users who download or connect to these resources may currently use them in the context of our approach by modifying our code at <https://github.com/SiminaB/CDGnet>. Expert-curated precision oncology databases include Clinical Interpretations of Variants in Cancer (CIViC)<sup>5</sup>, Cancer Genome Interpreter<sup>37</sup>, OncoKB<sup>38</sup>, DEPO<sup>39</sup>, and PMKB<sup>40</sup>, while more general resources include ClinVar<sup>41</sup>. These additional sources may further strengthen the clinical annotations and evidence related to germline and somatic alterations in our database and provide options in between curated drug label information and DrugBank targets. CIViC (<https://civicdb.org/>) is an open access, open source, community-driven web resource that allows clinical interpretations of mutations related to cancer. Cancer Genome Interpreter (<https://www.cancergenomeinterpreter.org/biomarkers>) is an online tool that connects genes and drugs along with their effect and publication sources, not as a network format, but as a tabular format. OncoKB (<https://oncokb.org/>) is another online precision oncology knowledge base that contains information about the effects and treatment implications of specific cancer gene alterations. DEPO (<http://depo-dinglab.ddns.net/>) contains druggable variant information such as drug therapy, evidence levels (FDA-approved, Clinical Trials, Case Reports, Preclinical), and the cancer types for intended treatments. The Precision Medicine Knowledgebase (PMKB) (<https://pmkb.weill.cornell.edu/>) provides information about

clinical cancer variants and interpretations. We also note that we are using a simplified model for incorporating pathway information via the consideration of targets which are downstream of oncogenes – there are scenarios which we do not capture where upstream targeting can also be useful, for example in the case of a feedback loop.<sup>42–44</sup> We will incorporate more complex information in further iterations of our tool.

Consortia such as the Clinical Genome (ClinGen) resource's Somatic Cancer Working Group<sup>3,45</sup> and the Global Alliance for Genomic in Health's (GA4GH) Variant Interpretation for Cancer Consortium (VICC)<sup>46</sup> have ongoing efforts to standardize and harmonize the expert-curated data in these different knowledge bases, with the goal to enhance the interoperability between these databases. We will align the future development of CDGnet with the guidelines and consensus frameworks developed by these consortiums. CDGnet can also serve as an informative tool for oncologists, molecular pathologists and genomic scientists who routinely participate in Molecular Tumor Board discussions.

Tools similar to CDGnet include PreMedKB<sup>47</sup> and Drug Gene Interaction Network. PreMedKB (<http://www.fudan-pgx.org/premedkb/>) is an integrated precision medicine knowledgebase for interpreting relationships between diseases, genes, variants, and drugs. Drug Gene Interaction Network (<http://seqome.com/drug-gene-network/>) is a commercial tool offered by Seqome Inc that builds drug-gene interaction networks to predict clinical response from multi-omics datasets. The advantage of CDGnet over these tools is that our approach allows users to input specific alterations found in a patient's tumor and cancer type, and outputs therapy options ordered based on priority. Such a personalized tool may eventually expand the range of options of targeted therapies for cancer patients in a clinical setting, a key goal of precision oncology.

## Supplementary Information

**Supplementary Table 1.** Curated FDA-targeted cancer therapies using label information, last updated on July 12, 2018.

## Acknowledgments

Funding for this work was provided through NIH grants R21CA220398 and supplement R21CA220398-02, U41HG009649, and P30CA051008 (via pilot award.)

## References

1. Amado, R. G. *et al.* Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J. Clin. Oncol.* **26**, 1626–1634 (2008).
2. Waks, A. G. & Winer, E. P. Breast Cancer Treatment: A Review. *JAMA* **321**, 288–300 (2019).
3. Ritter, D. I. *et al.* Somatic cancer variant curation and harmonization through consensus minimum variant level data. *Genome Med.* **8**, 117 (2016).
4. Dolman, L. *et al.* ClinGen advancing genomic data-sharing standards as a GA4GH driver project. *Hum. Mutat.* **39**, 1686–1689 (2018).
5. Griffith, M. *et al.* CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat. Genet.* **49**, 170–174 (2017).
6. Boca, S. M., Panagiotou, O. A., Rao, S., McGarvey, P. B. & Madhavan, S. Future of Evidence Synthesis in Precision Oncology: Between Systematic Reviews and Biocuration. *JCO Precision Oncology* **1** (2018).
7. Rao, S. *et al.* Quantification and expert evaluation of evidence for chemopredictive biomarkers to personalize cancer treatment. *Oncotarget* **8**, 37923–37934 (2017).
8. Chang, W. *et al.* Shiny: web application framework for R. *R package version 0.11.1*, 106 (2015).
9. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat. Med.* **10**, 789–799 (2004).
10. Network, T. C. G. A. & The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
11. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).
12. US National Library of Medicine, National Institutes of Health. DailyMed. *DailyMed* (2014). Available at: <https://dailymed.nlm.nih.gov/dailymed/>. (Accessed: 12th July 2018)
13. Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).
14. Bailey, M. H. *et al.* Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **174**, 1034–1035 (2018).

15. Targeted Cancer Therapies. *US National Cancer Institute, National Institutes of Health* (2014). Available at: <https://www.cancer.gov/about-cancer/treatment/types/targeted-therapies/targeted-therapies-fact-sheet>. (Accessed: 12th July 2018)
16. U.S. Food and Drug Administration. In Vitro Diagnostics - List of Cleared or Approved Companion Diagnostic Devices (In Vitro and Imaging Tools). Available at: <https://www.fda.gov/medicaldevices/productsandmedicalprocedures/invitrodiagnostics/ucm301431.htm>. (Accessed: 12th July 2018)
17. US Food and Drug Administration. In Vitro Diagnostics - Nucleic Acid Based Tests. Available at: <https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/%20InVitroDiagnostics/ucm330711.htm>. (Accessed: 12th July 2018)
18. Tenenbaum, D. KEGGREST: Client-side REST access to KEGG. *R package version 1.22.01*, (2018).
19. Zhang, J. D. & Wiemann, S. KEGGgraph: a graph approach to KEGG PATHWAY in R and bioconductor. *Bioinformatics* **25**, 1470–1471 (2009).
20. Carlson, M. org. Hs. eg. db: Genome wide annotation for Human. *R package version 3.7.0*, (2018).
21. Thurnherr, T., Singer, F., Stekhoven, D. J. & Beerewinkel, N. Genomic variant annotation workflow for clinical applications. *F1000Res.* **5**, 1963 (2016).
22. Wagner, A. H. *et al.* DGIdb 2.0: mining clinically relevant drug–gene interactions. *Nucleic Acids Res.* **44**, D1036–D1044 (2016).
23. US Food and Drug Administration. Drug Approvals and Databases - Drugs@FDA Data Files. Available at: <http://www.fda.gov/Drugs/InformationOnDrugs/ucm079750.htm>. (Accessed: 12th July 2018)
24. Kim, S. *et al.* PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* **47**, D1102–D1109 (2019).
25. Bostock, M., Ogievetsky, V. & Heer, J. D<sup>3</sup>: Data-Driven Documents. *IEEE Trans. Vis. Comput. Graph.* **17**, 2301–2309 (2011).
26. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, I1 (2013).

27. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
28. Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163**, 506–519 (2015).
29. Formisano, L. *et al.* Association of FGFR1 with ER $\alpha$  Maintains Ligand-Independent ER Transcription and Mediates Resistance to Estrogen Deprivation in ER+ Breast Cancer. *Clin. Cancer Res.* **23**, 6138–6150 (2017).
30. Formisano, L. *et al.* Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. *Nat. Commun.* **10**, 1373 (2019).
31. Perez-Garcia, J., Muñoz-Couselo, E., Soberino, J., Racca, F. & Cortes, J. Targeting FGFR pathway in breast cancer. *Breast* **37**, 126–133 (2018).
32. Li, M. M. *et al.* Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J. Mol. Diagn.* **19**, 4–23 (2017).
33. Zhang, W., Chien, J., Yong, J. & Kuang, R. Network-based machine learning and graph theory algorithms for precision oncology. *NPJ Precis Oncol* **1**, 25 (2017).
34. Jiang, X.-L., Martinez-Ledesma, E. & Morcos, F. Revealing protein networks and gene-drug connectivity in cancer from direct information. *Sci. Rep.* **7**, 3739 (2017).
35. Zhang, Y. & Tao, C. Network Analysis of Cancer-focused Association Network Reveals Distinct Network Association Patterns. *Cancer Inform.* **13**, 45–51 (2014).
36. Cheng, F. *et al.* Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput. Biol.* **8**, e1002503 (2012).
37. Tamborero, D. *et al.* Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* **10**, 25 (2018).
38. Chakravarty, D. *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol* **2017**, (2017).
39. Sun, S. Q. *et al.* Database of evidence for precision oncology portal. *Bioinformatics* **34**, 4315–4317

- (2018).
40. Huang, L. *et al.* The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. *J. Am. Med. Inform. Assoc.* **24**, 513–519 (2017).
  41. Landrum, M. J. *et al.* ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, D862–8 (2016).
  42. Prahallad, A. *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* **483**, 100–103 (2012).
  43. Solit, D. B. & Jänne, P. A. Translational medicine: Primed for resistance. *Nature* **483**, 44–45 (2012).
  44. Chandarlapaty, S. *et al.* AKT Inhibition Relieves Feedback Suppression of Receptor Tyrosine Kinase Expression and Activity. *Cancer Cell* **19**, 58–71 (2011).
  45. Madhavan, S. *et al.* ClinGen Cancer Somatic Working Group - standardizing and democratizing access to cancer molecular diagnostic data to drive translational research. *Pac. Symp. Biocomput.* **23**, 247–258 (2018).
  46. Wagner, A. H. *et al.* A harmonized meta-knowledgebase of clinical interpretations of cancer genomic variants. *bioRxiv* 366856 (2018). doi:10.1101/366856
  47. Yu, Y. *et al.* PreMedKB: an integrated precision medicine knowledgebase for interpreting relationships between diseases, genes, variants and drugs. *Nucleic Acids Res.* **47**, D1090–D1101 (2019).