Background:
Next generation sequencing of cancer genomes has revolutionized the field of precision oncology in recent years. Using this technology, it is now possible to classify cancer subtypes based on the similarity of their molecular profiles. With biomarkers-driven clinical trials such as NCI-MATCH [1], it is now possible to treat specific genomic profiles of tumors regardless of their cancer type. This revolution of genomic-based therapeutics has advanced the march towards legitimate precision oncology. However, there is a need for researchers to be able to have a resource to query molecular profiles and connect those profiles to approved or investigational therapeutics.

Here we present IMPACT Web Portal, a database linking the molecular profiles of tumors to clinical and pre-clinical oncology actionable therapeutics. We previously published a whole exome sequencing (WES) analysis pipeline, IMPACT (Integrating Molecular Profiles with ACtionable Therapeutics) that matches molecular profiles with actionable therapeutics [2]. However, this tool is currently only available for WES data and requires command line-level programming.
skills. To facilitate the translational ability of this method, we developed a web-based database that requires no programming skills and is applicable to any type of sequencing data source.

Several databases have been developed to provide drug-gene interactions [3–6], some of the databases are focusing in cancer with drug-target gene variants information [7–10]. The IMPACT Web Portal differs from existing resources in the following aspects: (i) IMPACT web portal includes actionable therapeutics integrated from ten different data sources and includes all approved oncology drugs, a variety of current investigational drugs in cancer clinical trials, and pharmacogenetics databases. (ii) IMPACT Web Portal allows for users to input molecular profiles of individual tumors. The search can include genes, variants, fusions, and copy number changes that will each link to all known actionable therapeutics in the database. (iii) IMPACT Web Portal uses a drug-based database to rank potentially therapeutic compounds into 3 levels: Level 1 contains all approved drugs with variant-level (Level 1A) and gene-level (Level 1B) evidence. Level 2 contains drugs currently in cancer clinical trials. Level 3 uses pharmacogenetics to link altered gene targets to potentially actionable therapeutics. (iv) A hypergeometric test is used to calculate a \( p \)-value in order to rank each drug by its specificity to the molecular profile. (v) IMPACT Web Portal links information to other resources for continued investigation of drug-gene interactions. Each drug name is a link, taking users to a drug-oriented results page listing other gene targets of the drug, and when available the structure and PubChem identification number of the drug. Each gene also links to the external NCBI gene database. (iv) IMPACT Web Portal has the largest collection of genes, variants, fusions, and copy number changes, linked to the largest number of actionable therapeutics when compared to other oncology databases. Here, we describe the IMPACT Web Portal, an online, user-friendly, database that connects a tumor’s molecular profile to actionable therapeutics integrated from ten of the most well curated data sources. We also provide an example illustrating the utility of the IMPACT Web Portal.

**Construction and content**

**IMPACT database construction**

Figure 1 illustrates the development workflow of the IMPACT database. We extracted drug-target genes
and variants from ten data sources: the Drug Repurposing Hub [3], Food and Drug Administration (FDA) website (FDA.gov) [11], A Comprehensive Map of Molecular Drug Targets of FDA-approved drugs published in Nature Reviews Drug Discovery [4], DSigDB [6], DGIdb [5], OncoKB [7], My Cancer Genome [8], the MD Anderson Precision Cancer Therapy database [10], drug-gene relationships in the National Cancer Institute (NCI) MATCH clinical trials [1, 9], and Clinical Implementation of Pharmacogenetics Consortium (CPIC) [12]. Table 1 provides the descriptions of these data sources. We retrieved all the synonyms and International Chemical Identifier (InChI) or InChIKey of the compiled compounds list from PubChem [13]. We used InChI and InChIKey to identify and unify compounds in the list. To unify target genes and proteins, we used UniProt [14] to convert protein names to NCBI Entrez Gene Symbols [15]. For drugs that target gene fusions, we extracted additional known gene fusions from ChimerDB [16] and the Tumor Fusion Gene Data Portal [17]. We then queried the drugs list against ClinicalTrials.gov to retrieve all drugs tested in cancer clinical trials. Through these steps, we collected 776 drugs, 1326 target genes and 435 target gene variants. We developed the IMPACT database using MySQL version 14.14 (Distribute 5.7.11) on the OSX 10.11 (x86_64) platform. We used Python Version 2.7.11 to write scripts to perform data wrangling.

**IMPACT database content**

We classified the drugs and compounds collected in the IMPACT database into three levels, based on the level of evidence of drug-target genes (Fig. 2).

**Level 1 approved drugs**

Level 1 contains 221 approved oncology drugs as of Sept 20, 2017. Level 1 is further divided into two sub levels: Level 1A comprises approved drugs with approved drug-target gene variants (including mutations, amplification, deletion and fusions); Level 1B consists of approved drugs with known target genes. Level 1A is composed of 47 approved drugs targeting 47 genes and 265 gene variants, fusions or copy number changes. Level 1B contains 221 oncological approved drugs that target 1170 genes.

**Level 2 investigational therapeutics**

Level 2 consists of 390 drugs currently being tested in cancer clinical trials (as of Sept 20, 2017). This set of investigational therapeutics target 370 genes.

**Level 3 pharmacogenetics**

Level 3 contains 203 approved drugs and their interactions with 110 genes obtained from the Clinical Pharmacogenetics Implementation Consortium (CPIC).

---

**Table 1 Data Sources used in IMPACT Web Portal Database Construction**

| Data Source | Description | Reference |
|-------------|-------------|-----------|
| OncoKB      | Precision oncology knowledge base of 100 drugs and their 476 genes and 3753 variants. | [7] |
| CPIC        | Pharmacogenetic knowledge of drug-gene interactions of 203 drugs. | [12] |
| MD Anderson Precision Medicine | Contains 416 approved and investigational oncological drugs and their target genes. | [10] |
| My Cancer Genome | Contains 237 approved and investigational oncological drugs and their target genes. | [8] |
| Nature Reviews Drug Discovery: A comprehensive map of molecular drug targets | A comprehensive mapping of 1578 US FDA-approved drugs and their 667 human targets. | [4] |
| Food and Drug Administration (FDA.gov) | Contains drug labelling of genes and variants for approved drugs. | [11] |
| The Drug Repurposing Hub | Extensive annotations of drugs and the genes they target for drug repurposing research. The current version contains 5628 compounds targeting 2172 proteins. | [3] |
| DSigDB      | Approved and investigational therapeutics of drug gene signatures collected from PubChem/ChEMBL and kinase inhibition experiments. | [6] |
| DGIdb       | Drug-gene interactions database collected from ten databases and 41 gene categories. | [5] |
| NCI-MATCH Trial | Drugs and associated target genes used to recruit various cancer patients (ClinicalTrials.gov NCT02465060). | [1, 9] |
| Tumor Gene Fusion Data Portal | Data base contains 8695 gene fusions detected from the Cancer Genome Atlas RNA-sequencing data. | [17] |
| ChimerDB    | Comprehensive database of 1066 gene fusions encompassing analysis of RNA-sequencing data, PubMed Abstract text mining and manual curations. | [16] |
IMPACT web portal

We developed a web portal and user interface to query IMPACT database using JavaScript and jQuery (Fig. 3a). The IMPACT Web Portal allows users to query, search, view and download data. In the query box, users are required to enter at least one gene (official gene symbol) followed by an optional alteration which can include a variant, copy number change, fusion, or fusion partner. For example, a user may enter any or all of the following on separate lines: BRAF, BRAF(V600E), BRAF V600E, BRAF(AMP), BRAF(fusion). Each entry will be queried in the database and mapped to potential actionable therapeutics in the IMPACT database. A hypergeometric test is conducted to compute a \( p \)-value for each mapped drug-target genes and drug-target gene variants.

Results from the query were returned in the IMPACT results page (Fig. 3b). The results page is divided into three levels based on the level of evidence. Within each level, each drug is listed followed by the mapped actionable genes and variants. Mapped drugs within each level were sorted by \( p \)-value (Fig. 3b). We developed a compound page to provide additional information for the drug (Fig. 3c). For the top part of the compound page, we used RDKit to generate the molecular descriptors for the compound. We used Marvin Sketch to draw the molecular structure. External links to PubChem is also provided in the compound page. The middle part of the compound page provides the target genes and variants of the drug, as well as the data source of the drug-target gene interactions. The bottom part of the compound page provides the list of Clinical Trials investigated by the compound.

Data availability

IMPACT Web Portal is freely available for non-commercial research only use at http://tanlab.ucdenver.edu/IMPACT. IMPACT Web Portal data is available to download as tab-delimited plain text (.txt) files.

Utility and discussion

To illustrate the utility of the IMPACT Web Portal, we performed a query from the analysis of whole-exome sequencing (WES) data. Previously, we published our
The IMPACT Web Portal is freely available to all users at http://tanlab.ucdenver.edu/IMPACT. This web portal is accessible by web browser.

### Conclusions

In conclusion, we developed IMPACT Web Portal, a novel online database for connecting molecular profiles to actionable therapeutics. The IMPACT Web Portal online resource allows users to search and connect 1326 target genes and 435 target gene variants against 776 approved and investigational cancer drugs. By utilizing three distinct levels of actionable therapeutics, users are able to find drugs already approved (Level 1), currently being tested in clinical trials (Level 2), and with pharmacogenetic evidence (Level 3). We believe that IMPACT Web Portal represents a significant improvement in the ability to connect molecular profiles with actionable therapeutics by using up to date resources. The user-friendly IMPACT Web Portal allows users to search for molecular profiles of individual tumors from any sequencing data and match them to actionable therapeutics for translational or drug repurposing oncology studies.

### Availability and requirements

The IMPACT Web Portal is freely available to all users at [http://tanlab.ucdenver.edu/IMPACT](http://tanlab.ucdenver.edu/IMPACT). This web portal is accessible by web browser.

### Abbreviations

- CPIC: Clinical Implementation of Pharmacogenetics Consortium
- FDA: Food and Drug Administration
- IMPACT: Integrating Molecular Profiles with Actionable Therapeutics
- InChI: International Chemical Identifier
- NCI: National Cancer Institute
- WES: Whole Exome Sequencing

### Acknowledgements

We thank the Tan Lab members and the reviewers for their useful comments on this manuscript.

### Funding

The publication cost of this article was partly funded by the Cancer League of Colorado, the National Institutes of Health (P30CA46934, P50CA058187), the Amy Davis Foundation, the Moore Family Foundation, and the David F. and Margaret T. Grohne Family Foundation.

### Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the IMPACT WEB PORTAL repository, [http://tanlab.ucdenver.edu/IMPACT](http://tanlab.ucdenver.edu/IMPACT).

### About this supplement

This article has been published as part of BMC Medical Genomics Volume 11 Supplement 2, 2018: Proceedings of the 28th International Conference on Genome Informatics: medical genomics. The full contents of the supplement are available online at [https://bmcmedgenomics.biomedcentral.com/articles-supplements/volume-11-supplement-2](https://bmcmedgenomics.biomedcentral.com/articles-supplements/volume-11-supplement-2).

### Authors’ contributions

JDH wrote the initial draft of the manuscript, design of the user interface, collecting the data and institutional affiliations. ACT conceived the idea, design the user interface, feedback and tested the database. MY built the database and testing of the biological case study. JDH wrote the initial draft of the manuscript, design of the user interface, feedback and tested the database. ACT conceived the idea, design the user interface, implemented the biological use example, created the figures and wrote the final manuscript. All authors read and approved the final manuscript.

### Ethics approval and consent to participate

Tissue acquisition from consenting melanoma patients at the time of removal of a primary tumor or biopsy was conducted under a Colorado Multi-Institutional Review Board–approved protocol, COMIRB-05-0309.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

1Division of Medical Oncology, Department of Medicine, Translational Bioinformatics and Cancer Systems Biology Laboratory, University of Colorado Anschutz Medical Campus, Aurora, CO, USA
2Division of Medical Oncology, Department of Medicine, Robinson Melanoma Laboratory, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

Published: 20 April 2018

### References

1. McNeil C. NCI-MATCH launch highlights new trial design in precision-medicine era. J Natl Cancer Inst. 2015;107(7):djv193.
2. Hintzsche J, Kim J, Yadav V, Amato C, Robinson SE, Seelenfreund E, Shellman Y, Wisell J, Applegate A, McCarter M, et al. IMPACT: a whole-exome sequencing analysis pipeline for integrating molecular profiles with actionable therapeutics in clinical samples. J Am Med Inform Assoc. 2016;23(4):721–30.
3. Corsello SM, Bittker JA, Liu Z, Gould J, McCaren P, Hirschman JE, Johnston SE, Vticc A, Wong B, Khan M, et al. The drug repurposing hubs: a next-generation drug library and information resource. Nat Med. 2017;23(4):405–8.
4. Santos R, Ursu O, Gauton A, Bento AP, Donadi RS, Bologna CG, Karlsson A, Al-Lazikani B, Hersey A, Oprea TI, et al. A comprehensive map of molecular drug targets. Nat Rev Drug Discov. 2017;16(1):19–34.

5. Wagner AH, Coffman AC, Ainscough BJ, Spies NC, Skidmore ZL, Campbell KM, Kysiaik K, Pan D, McMichael JF, Eldred JM, et al. DGIdb 2.0: mining clinically relevant drug-gene interactions. Nucleic Acids Res. 2016;44(D1):D1036–44.

6. Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S, Jeon M, Kang J, Tan AC. DSigDB: drug signatures database for gene set analysis. Bioinformatics. 2015;31(18):3069–71.

7. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, Rudolph JE, Yaeger R, Soumerai T, Nisan MH, et al. OncoKB: a precision oncology Knowledge Base. JCO Precis Oncol. 2017;2017 https://doi.org/10.1200/PO.17.00011.

8. Michiel CM, Lovly CM, Levy MA. My Cancer genome. Cancer Genetics. 2014;207(6):289.

9. Conley BA, Doroshow JH. Molecular analysis for therapy choice: NCI MATCH. Semin Oncol. 2014;41(3):297–9.

10. Kurnit KC, Bailey AM, Zeng J, Johnson AM, Shufean MA, Busco L, Lützenburger BC, Sánchez NS, Khotskaya YB, Holla V, Simpson A. Personalized Cancer therapy: a publicly available precision oncology resource. Cancer Res. 2017;77(21):e123–6.

11. US Food and Drug Administration. Available at: http://www.fda.gov. Accessed 18 Dec 2017.

12. Luzum JA, Palayc RE, Elsey AR, Haidar CE, Peterson JF, Whirl-Carrillo M, Handelman SK, Palmer K, Pulley JM, Beller M, et al. The pharmacogenomics research network translational pharmacogenetics program: outcomes and metrics of Pharmacogenetic implementations across diverse healthcare systems. Clin Pharmacol Ther. 2017;102(3):502–10.

13. Wang Y, Bryant SH, Cheng T, Wang J, Gindulyte A, Shoemaker BA, Thiessen PA, He S, Zhang J. PubChem BioAssay: 2017 update. Nucleic Acids Res. 2017;45(D1):D955–63.

14. The UniProt Consortium. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 2017;45(D1):D158–69.

15. NCBI. Resource coordinators: database resources of the National Center for biotechnology information. Nucleic Acids Res. 2016;44(D1):D7–19.

16. Lee M, Lee K, Yu N, Jang I, Choi I, Kim P, Jang YE, Kim B, Kim S, Lee B, et al. ChimerDB 3.0: an enhanced database for fusion genes from cancer transcriptome and literature data mining. Nucleic Acids Res. 2017;45(D1):D784–9.

17. Yoshihara K, Wang Q, Torres-Garcia W, Zheng S, Vegesna R, Kim H, Verhaak RG. The landscape and therapeutic relevance of cancer-associated transcript fusions. Oncogene. 2015;34(37):4845–54.

18. Kwong LN, Costello JC, Liu H, Jiang S, Helms TL, Langsdorf AE, Jakubosky D, Genovese G, Muller FL, Jeong JH, et al. Oncogenic NRAS signaling differentially regulates survival and proliferation in melanoma. Nat Med. 2012;18(10):1503–10.