

Q-omics: Smart Software for Assisting Oncology and Cancer Research

Jieun Lee^{1,3}, Youngju Kim^{1,3}, Seonghee Jin¹, Heeseung Yoo¹, Sumin Jeong¹, Euna Jeong², and Sukjoon Yoon^{1,2,*}

¹Department of Biological Sciences, Sookmyung Women's University, Seoul 04310, Korea, ²Research Institute of Women's Health, Sookmyung Women's University, Seoul 04310, Korea, ³These authors contributed equally to this work.

*Correspondence: yoonsj@sookmyung.ac.kr
<https://doi.org/10.14348/molcells.2021.0169>
www.molcells.org

The rapid increase in collateral omics and phenotypic data has enabled data-driven studies for the fast discovery of cancer targets and biomarkers. Thus, it is necessary to develop convenient tools for general oncologists and cancer scientists to carry out customized data mining without computational expertise. For this purpose, we developed innovative software that enables user-driven analyses assisted by knowledge-based smart systems. Publicly available data on mutations, gene expression, patient survival, immune score, drug screening and RNAi screening were integrated from the TCGA, GDSC, CCLE, NCI, and DepMap databases. The optimal selection of samples and other filtering options were guided by the smart function of the software for data mining and visualization on Kaplan-Meier plots, box plots and scatter plots of publication quality. We implemented unique algorithms for both data mining and visualization, thus simplifying and accelerating user-driven discovery activities on large multiomics datasets. The present Q-omics software program (v0.95) is available at <http://qomics.sookmyung.ac.kr>.

Keywords: biomarker, cancer bioinformatics, immune infiltrate, Kaplan-Meier plot, omics data mining, smart software

INTRODUCTION

Large collateral datasets, including those on mutations, gene expression, drug/RNAi screening and patient survival, are publicly available from diverse resources (Barretina et al., 2012; Cancer Genome Atlas Research Network et al., 2013; Ghandi et al., 2019; Guan et al., 2019; Iorio et al., 2016; Monks et al., 2018; Shi et al., 2021). Integrated analysis of the cross-association of these datasets provides useful clues for finding novel targets, predictive biomarkers and related mechanisms (Jeong et al., 2020; Shen et al., 2019). For example, many genes and mutations have been found to be associated with the patient survival rate via analyses of datasets from the TCGA database (Cao et al., 2020; Eckstein et al., 2020; Hong et al., 2017; Kitsou et al., 2020; Yang et al., 2011; Zhong et al., 2020). Cell line databases provide clues for the identification of predictive biomarkers against drug resistance and/or sensitivity (Garnett et al., 2012; He et al., 2014; Kim et al., 2016; Li et al., 2021; Yang et al., 2013). Novel targets against subtype-specific cancer mutations have also been suggested (Biswas et al., 2019; Li et al., 2019; Park et al., 2019).

An explosive increase in these collateral datasets will provide important resources for diverse data-driven cancer research projects. However, systematic and integrated analyses of these datasets are still challenging to most oncologists and cancer researchers with no computational background. Many

Received 25 June, 2021; revised 22 August, 2021; accepted 4 September, 2021; published online 17 November, 2021

eISSN: 0219-1032

©The Korean Society for Molecular and Cellular Biology.

©This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>.

web-based tools have been developed to improve the utility of public cancer datasets, such as OncoPrint (Rhodes et al., 2004), cBioPortal (Cerami et al., 2012), and TIMER2.0 (Li et al., 2020). Although these web-based applications provide useful tools for a quick data search with significant information, user-oriented customized calculation and data filtration are generally limited from these server-provided functions. Thus, flexible and comprehensive software is required for cancer scientists to carry out customized data processing and computation on their local computers.

Here, we attempted to develop innovative smart software for oncologists to easily start their own data mining projects without computational skills. We established two aims for this software. First, the process of data analysis and visualization should be simple and comprehensive by providing a user-friendly graphical interface and an intuitive organization of menus. Second, we tried to implement smart functions that guide users to find optimal outputs, i.e., associated data pairs and graphs, via real-time communication with a server-side knowledge base harboring billions of pre-calculated data pairs. For these purposes, we simultaneously developed stand-alone software with data processing and computation abilities and a server-side knowledge base that can be connected to local software. This report briefly presents the functions and utilities of this software, Q-omics v0.95. The smart system of the implemented knowledge base will be continuously updated with improved visualization options in the user interface. We expect that the present computer-aided, smart data mining system will have general utilities in all fields of oncology and cancer research without the requirements of bioinformatics skills.

MATERIALS AND METHODS

Cell line data

Cell line-based large-scale data consisting of RNA sequencing data (Expression, ver. 20Q1), sgRNA sequencing data (CRISPR, ver. 20Q1), shRNA screening data (Achilles + DRIVE + Marcotte, DEMETER2), mutation data (Mutation Public, ver. 20Q4), and drug response data (Sanger GDSC1 and GDSC2) were obtained from the DepMap portal (<https://depmap.org/>

portal/). RNA sequencing data represent \log_2 -transformed transcripts per million (TMP) + 1 values using RSEM normalization. sgRNA and shRNA data are batch-corrected CERES gene knockout effects (Meyers et al., 2017) and DEMETER2 estimated gene knockdown effects (McFarland et al., 2018), respectively. Mutation data are MAF of gene mutations. Drug response data are published as IC50 (nM) values and we transformed to logarithmic scale pIC50 (M). To analyze associations between datasets, 20 lineages with a sufficient number of common cell lines between RNA sequencing data and other data (sgRNA and drug response) were used in this study. Furthermore, the gene expression data of NCI60 cell lines treated with 15 drugs were obtained from the GEO database (GSE116436) (Monks et al., 2018). Details on the cell line, number of lineages, number of cell lines, and number of genes/drugs are shown in Table 1.

Tissue data

Patient RNA sequencing data, clinical data, and mutation data were obtained from the Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>). In total, 33 cancer types were investigated in this study. For comparisons between normal and tumor data, paired normal and tumor tissue samples from 18 cancer types whose number of matched tissue samples was larger than 2 were collected. RNA sequencing data in FPKM (fragments per kilobase of transcript per million fragments mapped) values are transformed to \log_2 (TPM + 1) values after downloading. In addition, immune cell enrichment score for TCGA data was obtained from the xCell portal (<https://xcell.ucsf.edu/>) (Aran et al., 2017). Details on the tissue type, number of lineages, number of samples, and number of genes are shown in Table 1.

Cross-association analysis

To analyze associations between two datasets, we performed a cross-association analysis between phenotypic efficacy and gene expression in our previous work (Jeong et al., 2020). In this study, we extended the concept of cross-association to analyze more diverse datasets, including those on gene expression, sgRNAs, shRNAs, the drug response, and mutations (Fig. 1). Cross-associations between each data type, such as

Table 1. Numbers of data points integrated into Q-omics software

	No. of lineages	No. of cell lines/ No. of samples	No. of genes/ No. of drugs	Data type
Cell line data				
Gene expression	20	1,061	19,137	RNA sequencing
sgRNA	20	741	18,110	CRISPR
shRNA	20	587	16,800	RNAi shRNA
Drug response	20	1,001	397	Drug response
Mutation	20	1,281	18,731	Exome sequencing
Drug-induced gene expression	13	60	12,305/15	DNA microarray
Tissue data				
Tumor gene expression	33	9,951	38,311	RNA sequencing
Paired normal vs. cancer: gene expression	18	679	38,311	RNA sequencing
Mutation	33	9,100	20,850	Exome sequencing
Immune	33	8,954	64 (cell types)	Cell type enrichment score

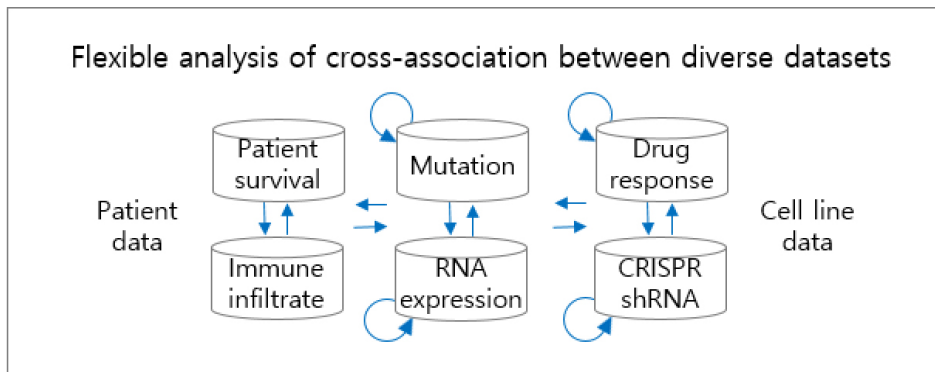


Fig. 1. Overview of data integration in Q-omics software. Public datasets from the TCGA, GDSC, CCLE, NCI and DepMap were integrated for the cross-association analysis (blue arrow) of between any two datasets.

drug versus RNA-seq and shRNA versus mutation, and within the same data type, such as sgRNA versus sgRNA and mutation versus mutation, can be analyzed.

Two association measures are predictivity and descriptivity. Given any two datasets (X and Y), we assume that x and y are entries of X and Y, respectively. The predictivity of x measures the difference in x values between two groups divided by the median of y . In contrast, the descriptivity of x measures the difference in y values between two groups divided by the median x . Significance was tested using Fisher's exact test for categorical data (mutation) and Student's t -test for numerical data (all other data).

Survival analysis

Survival data were analyzed using the Kaplan–Meier (KM) method, and the log-rank test was used to compare the survival outcomes of two groups as a test of statistical significance. Furthermore, the area under the curve (AUC) was calculated to provide an estimate of the size of the difference between two groups. In this study, overall survival (OS) and disease-free survival (DFS) were analyzed. The two analyses differed according to the definition of the primary endpoint: all causes of death during the study period were used to analyze OS, and a tumor event or death was used to analyze DFS.

For the single-gene survival analysis, patients were divided into two groups based on high or low expression of the given gene or mutation status of the given gene. The association of two genes can be determined in advance to generate several subgroups for combined-gene survival analysis. Furthermore, for a more sophisticated survival analysis, a subset of patients was selected using clinical information such as sex, stage, or any combination of sex and stage.

Smart search

Q-omics was designed to run locally on user computers. While running Q-omics, time-consuming or data/memory-intensive analyses are performed on the server computer. For example, the cross-association analysis on the user side investigates only in a given lineage, while the smart search retrieves the most highly associated pairs in all 20 lineages from the server side. Similarly, for the survival analysis, the user side calculates the survival rate based on a single gene in a given lineage, while the smart search provides the significance of the survival rate based on a given gene in all 33 lineages.

Box plot analysis

Box plots in Q-omics can be used to visualize differences in the distribution of numerical data between different groups. The differences between two groups were analyzed by calculating the fold change and P value (Student's t -test).

Q-omics also provides a platform for comparisons between drug-induced changes in gene expression. Gene expression data from NCI60 cell lines treated with 15 anticancer agents contained the measured expression values of nine genes at three time points (2, 4, and 24 h) and at three doses (0 nM, low dose, and high dose) (Monks et al., 2018). The low and high doses used varied depending on each drug. For these data, box plots were generated to compare time- and dose-dependent gene expression. Groups were divided based on time points or doses, and box plots were used to display fold changes between time points (4 h vs 2 h and 24 h vs 2 h) or between doses (low dose vs 0 nM and high dose vs 0 nM), respectively, not raw gene expression.

Scatter plot analysis

Scatter plots were used to display relationships between two numeric variables, and the strength and direction of the linear relationships were assessed by Pearson's correlation coefficient in Q-omics.

Q-omics implementation

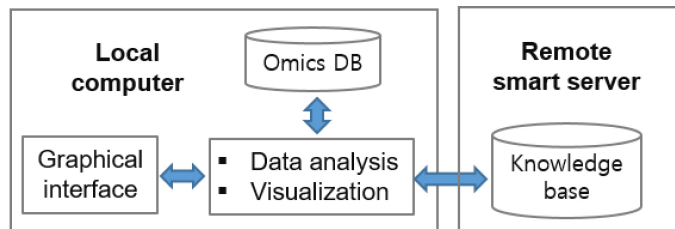
Q-omics was implemented in Python 3, and MySQL was used for the smart search.

RESULTS AND DISCUSSION

Q-omics software runs on the user's computer, providing a graphical interface and computational/visualization modules together with its own local database (Fig. 2A). To assist in user data mining, Q-omics interacts with a server-side knowledge base and retrieves relevant information for analysis. The knowledge base harbors billions of precalculated, significantly associated data pairs with related information such as sample filters and calculation options. Smart algorithms in the knowledge base promptly select data pairs and information that is relevant to the user's query and then returns it to Q-omics.

As described in Fig. 1, users can start data mining with one query (i.e., gene expression, mutations, drugs, or sh/sgRNAs). The front page of Q-omics provides a graphical interface for selecting the analysis type, query and sample

A



B

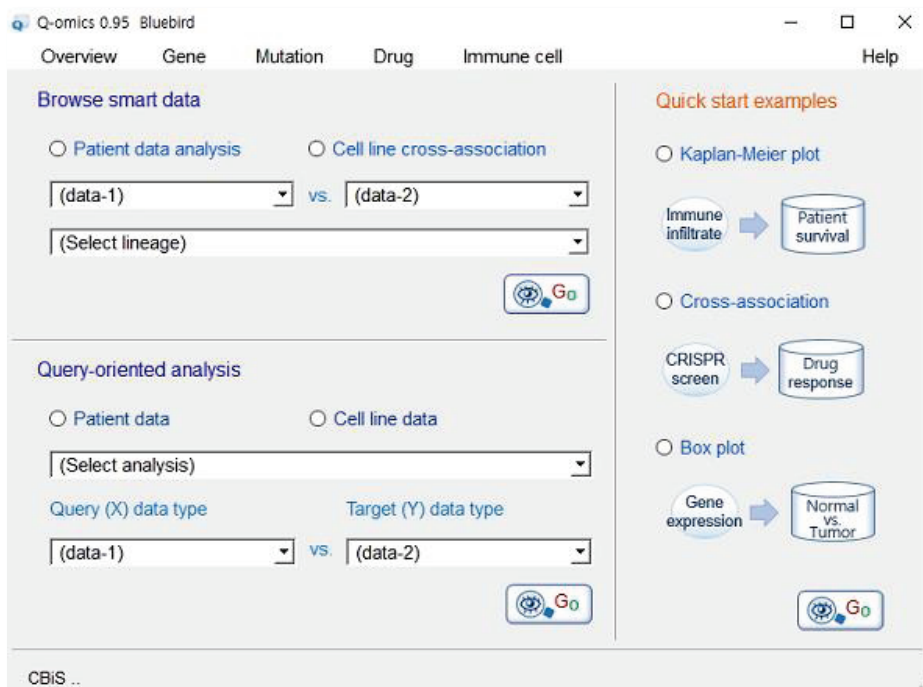


Fig. 2. Software workflow and user interface. (A) The workflow of functional modules and databases between the local software and server-side knowledge base in Q-omics. (B) Main interface of Q-omics software. Search options are separated into “Browse smart data” and “Query-oriented analysis”. “Quick start examples” are comprehensive options for first-time users. Knowledge-based smart search is enabled for all of the search options.

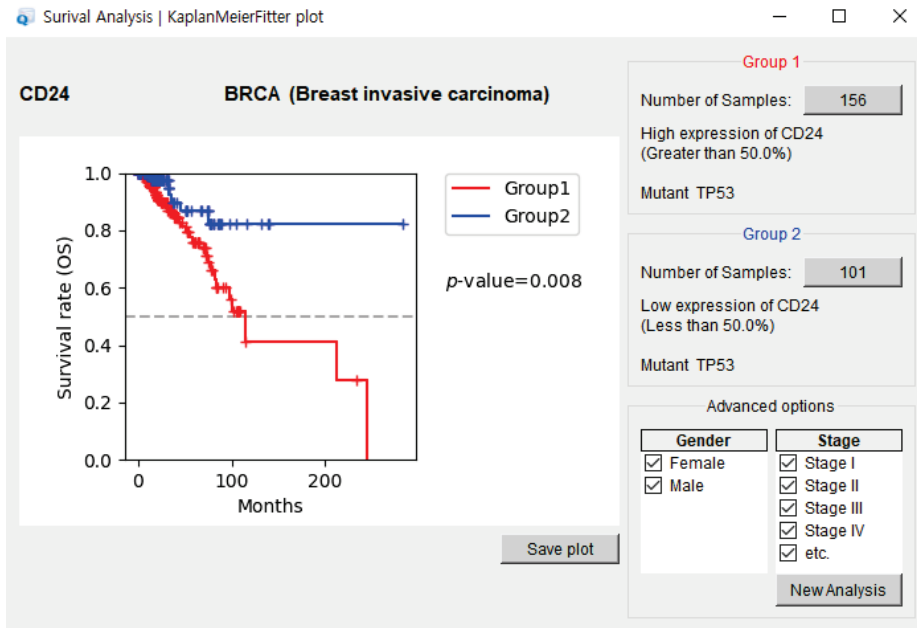
type (Fig. 2B). Basically, all analyses are separated into those with patient samples and those with cell lines. Available analyses with patient samples are as follows: (1) survival analyses (Kaplan–Meier plots) according to gene expression and mutations, (2) differential gene expression analyses between normal and cancer cells, and (3) scatter/box plots analyses of gene expression and/or mutation pairs. Available analyses with cell lines are as follows: (1) cross-association analyses between any pair of datasets according to gene expression, mutations, shRNA screening data, sgRNA screening data and drug screening data, (2) change (induction) analyses of gene expression before/after drug treatments, and (3) scatter/box plot analyses of pairs according to gene expression, mutations, shRNAs, sgRNAs and drugs. The menu “Quick start examples” is used to demonstrate graphical outputs and smart functions of the software using the preselected analysis type and user-selected queries. In all analyses, the resulting graphs and data can be saved for further usage.

Fig. 3A demonstrates the survival analysis module of the software. A Kaplan–Meier plot of BRCA patient data was generated by using user-selected options: CD24 gene expression with TP53 mutations. The graphical panel provides

detailed information on selected samples and further filtering options such as sex and stage. Together with the panel of Kaplan–Meier plots, Q-omics software provides a panel of smart search results (Fig. 3B). This smart panel provides a list of genes that exhibit significant ($P < 0.01$) associations with the survival rate in combination with user-selected queries, i.e., CD24 gene expression. Users can select one of the genes in the list and see the Kaplan–Meier plot in the new panel. This is very useful for the quick discovery of gene expression changes or mutations that are associated with the queried gene (user’s interest) in the patient survival analysis. This smart list is automatically generated from the server-side knowledge base by using information such as user-selected queries and lineages. The smart system in the server searches genes or mutations that are related (i.e., significantly associated) to the user’s interests from the knowledge base and sends them to the Q-omics user interface. Algorithms in the smart system are improved and updated continuously with the increase in data in the knowledge base.

Fig. 4A shows the Q-omics output panel of a cross-association analysis between the user-selected drug, cisplatin, and 17,795 sgRNAs in lung cancer cell lines. The present example

A



B

Survival Analysis | Smart search

Other significant options for your query

Gene expression. associated ($p < 0.01$) with CD24 expression *AUC: Area Under Curve

	Target	Measure	Exp. filter	Gender	Stage	Group1 AUC	Group2 AUC	p -value	Combi.
97	ZFP91P1	OS	Median	Both	All	0.426	0.71	0.002	+
98	POMT2	OS	Median	Both	All	0.738	0.407	0	+
99	GPR1-AS	OS	Median	Both	All	0.439	0.721	0	+
100	MIR7156	OS	Median	Both	All	0.49	0.732	0.004	+
101	ACSM3	OS	Median	Both	All	0.45	0.728	0	-
102	SOX8	OS	Median	Both	All	0.428	0.669	0	-
103	STAG3L5P-PVRIG2P-PILRB	OS	Median	Both	All	0.46	0.725	0	-
104	MPO	OS	Median	Both	All	0.431	0.716	0	-
105	RIMBP3B	OS	Median	Both	All	0.504	0.661	0.007	-

New KM plot

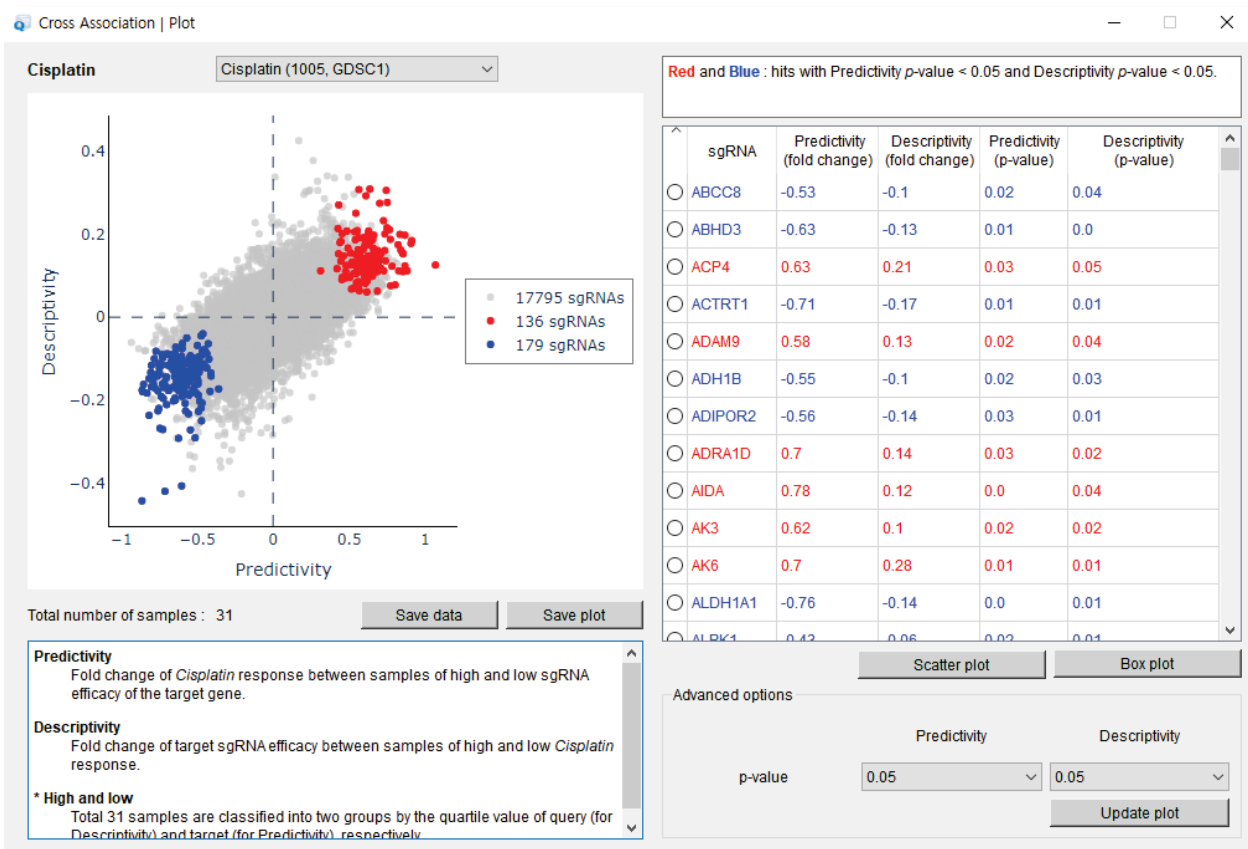
Fig. 3. Graphical interface of patient survival analysis and related smart search results. (A) The panel of survival analyses included Kaplan-Meier (KM) plots, sample group information and advanced options for plotting. (B) The panel of gene lists retrieved by the smart algorithm from the server-side knowledge base. In this example, the list shows genes that are significantly ($P < 0.01$) associated with the user's query in the KM plot.

shows that the responses of 136 sgRNAs exhibit a positive association ($P < 0.05$) with the cisplatin response (red circle in Fig. 4A), while 179 sgRNAs exhibit a negative association (blue circle in Fig. 4A) with the cisplatin response. A detailed list of hit sgRNAs is displayed on the right side of the panel. Hit selection can be optimized by changing the p -value cutoff or sample separation option (i.e., median or quartile). Specific association patterns between hit sgRNAs and cisplatin can be displayed as box plots or scatter plots (Figs. 4B and 4C).

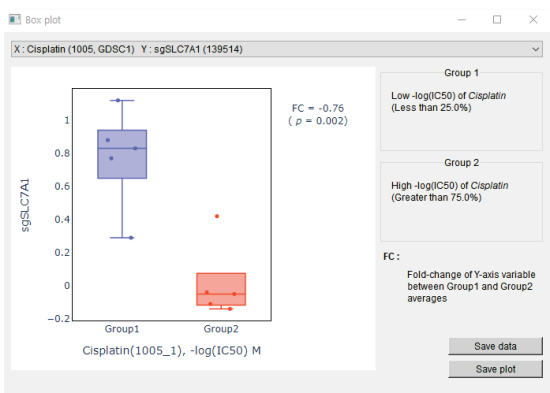
The predictivity and descriptivity measures from the

cross-association calculation were reported to be useful for the systematic evaluation of targets and biomarkers from multiomics data (Jeong et al., 2020). Q-omics software provides a simple and easy interface for calculating and analyzing the cross-association between any data pair, such as gene expression, mutations, sh/sgrNA screening data and drug screening data, from diverse resources. Q-omics also provides smart search results related to the user's query in the cross-association analysis. The software retrieves diverse association patterns with statistical significance to the user's query from

A



B



C

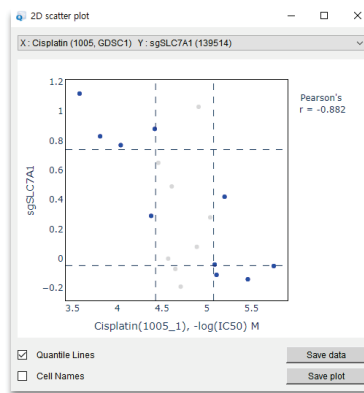


Fig. 4. Graphical interface of cross-association analysis between datasets using cell lines. (A) The panel of cross-associations displaying the predictivity and descriptivity scores of all data points. The list on the right side shows hits with significant P values. (B and C) Box plot and scatter plot of a selected hit from the cross-association panel. Box plots and scatter plots are also available for patient sample analyses.

the knowledge base and assists users in the optimal selection of data pairs and visualization.

In summary, Q-omics is an innovative software program

that enables users to carry out data mining and customized visualization without computational skills. The smart system of the software assists in the identification of new data pairs

related to/associated with the user's interests in real time. This software takes advantage of stand-alone software and web-based applications. Several discovery projects using this software are ongoing, and the results will be published in the near future.

ACKNOWLEDGMENTS

This work was financially supported by grants from the National Research Foundation of Korea (KRF), including the Science Research Center Program (NRF-2016R1A5A1011974), and the Mid-career Researcher Program (NRF-2017R1A2B2007745 and NRF-2018R1A2B6009313), funded by the Korean government (MEST).

AUTHOR CONTRIBUTIONS

S.Y. contributed to the overall study design. J.L., S.J. (Seonghee Jin), H.Y., S.J. (Sumin Jeong), E.J., and S.Y. conceived and implemented the software. J.L., Y.K., E.J., and S.Y. designed and implemented the database. J.L., Y.K., E.J., and S.Y. wrote manuscript.

CONFLICT OF INTEREST

The authors have no potential conflicts of interest to disclose.

ORCID

Jieun Lee <https://orcid.org/0000-0002-4554-965X>
 Youngju Kim <https://orcid.org/0000-0002-9176-1168>
 Seonghee Jin <https://orcid.org/0000-0002-0434-6215>
 Heeseung Yoo <https://orcid.org/0000-0002-4814-0315>
 Sumin Jeong <https://orcid.org/0000-0001-8193-8428>
 Euna Jeong <https://orcid.org/0000-0002-8802-4140>
 Sukjoon Yoon <https://orcid.org/0000-0001-7885-7117>

REFERENCES

Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* *18*, 220.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehar, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* *483*, 603-607.

Biswas, A., Haldane, A., Arnold, E., and Levy, R.M. (2019). Epistasis and entrenchment of drug resistance in HIV-1 subtype B. *Elife* *8*, e50524.

Cancer Genome Atlas Research Network, Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* *45*, 1113-1120.

Cao, R., Yuan, L., Ma, B., Wang, G., Qiu, W., and Tian, Y. (2020). An EMT-related gene signature for the prognosis of human bladder cancer. *J. Cell. Mol. Med.* *24*, 605-617.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* *2*, 401-404.

Eckstein, M., Strissel, P., Strick, R., Weyerer, V., Wirtz, R., Pfannstiel, C., Wullweber, A., Lange, F., Erben, P., Stoehr, R., et al. (2020). Cytotoxic T-cell-related gene expression signature predicts improved survival in muscle-invasive urothelial bladder cancer patients after radical cystectomy and adjuvant chemotherapy. *J. Immunother. Cancer* *8*, e000162.

Garnett, M.J., Edelman, E.J., Heidorn, S.J., Greenman, C.D., Dastur, A.,

Lau, K.W., Greninger, P., Thompson, I.R., Luo, X., Soares, J., et al. (2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* *483*, 570-575.

Ghandi, M., Huang, F.W., Jane-Valbuena, J., Kryukov, G.V., Lo, C.C., McDonald, E.R., 3rd, Barretina, J., Gelfand, E.T., Bielski, C.M., Li, H., et al. (2019). Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* *569*, 503-508.

Guan, N.N., Zhao, Y., Wang, C.C., Li, J.Q., Chen, X., and Piao, X. (2019). Anticancer drug response prediction in cell lines using weighted graph regularized matrix factorization. *Mol. Ther. Nucleic Acids* *17*, 164-174.

He, N., Kim, N., Song, M., Park, C., Kim, S., Park, E.Y., Yim, H.Y., Kim, K., Park, J.H., Kim, K.I., et al. (2014). Integrated analysis of transcriptomes of cancer cell lines and patient samples reveals STK11/LKB1-driven regulation of cAMP phosphodiesterase-4D. *Mol. Cancer Ther.* *13*, 2463-2473.

Hong, Y., Kim, N., Li, C., Jeong, E., and Yoon, S. (2017). Patient sample-oriented analysis of gene expression highlights extracellular signatures in breast cancer progression. *Biochem. Biophys. Res. Commun.* *487*, 307-312.

Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Goncalves, E., Barthorpe, S., Lightfoot, H., et al. (2016). A landscape of pharmacogenomic interactions in cancer. *Cell* *166*, 740-754.

Jeong, E., Lee, Y., Kim, Y., Lee, J., and Yoon, S. (2020). Analysis of cross-association between mRNA expression and RNAi efficacy for predictive target discovery in colon cancers. *Cancers (Basel)* *12*, 3091.

Kim, N., Yim, H.Y., He, N., Lee, C.J., Kim, J.H., Choi, J.S., Lee, H.S., Kim, S., Jeong, E., Song, M., et al. (2016). Cardiac glycosides display selective efficacy for STK11 mutant lung cancer. *Sci. Rep.* *6*, 29721.

Kitsou, M., Ayiomamitis, G.D., and Zaravinos, A. (2020). High expression of immune checkpoints is associated with the TIL load, mutation rate and patient survival in colorectal cancer. *Int. J. Oncol.* *57*, 237-248.

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., Li, B., and Liu, X.S. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* *48*(W1), W509-W514.

Li, W., Wang, H., Ma, Z., Zhang, J., Ou-Yang, W., Qi, Y., and Liu, J. (2019). Multi-omics analysis of microenvironment characteristics and immune escape mechanisms of hepatocellular carcinoma. *Front. Oncol.* *9*, 1019.

Li, Y., Umbach, D.M., Krahn, J.M., Shats, I., Li, X., and Li, L. (2021). Predicting tumor response to drugs based on gene-expression biomarkers of sensitivity learned from cancer cell lines. *BMC Genomics* *22*, 272.

McFarland, J.M., Ho, Z.V., Kugener, G., Dempster, J.M., Montgomery, P.G., Bryan, J.G., Krill-Burger, J.M., Green, T.M., Vazquez, F., Boehm, J.S., et al. (2018). Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. *Nat. Commun.* *9*, 4610.

Meyers, R.M., Bryan, J.G., McFarland, J.M., Weir, B.A., Sizemore, A.E., Xu, H., Dharia, N.V., Montgomery, P.G., Cowley, G.S., Pantel, S., et al. (2017). Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* *49*, 1779-1784.

Monks, A., Zhao, Y., Hose, C., Hamed, H., Krushkal, J., Fang, J., Sonkin, D., Palmisano, A., Polley, E.C., Fogli, L.K., et al. (2018). The NCI Transcriptional Pharmacodynamics Workbench: a tool to examine dynamic expression profiling of therapeutic response in the NCI-60 cell line panel. *Cancer Res.* *78*, 6807-6817.

Park, C., Lee, Y., Je, S., Chang, S., Kim, N., Jeong, E., and Yoon, S. (2019). Overexpression and selective anticancer efficacy of ENO3 in STK11 mutant lung cancers. *Mol. Cells* *42*, 804-809.

Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., and Chinnaiyan, A.M. (2004). ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* *6*, 1-6.

Shen, Y., Liu, J., Zhang, L., Dong, S., Zhang, J., Liu, Y., Zhou, H., and Dong, W. (2019). Identification of potential biomarkers and survival analysis for head and neck squamous cell carcinoma using bioinformatics strategy: a study based on TCGA and GEO datasets. *Biomed Res. Int.* 2019, 7376034.

Shi, B., Ding, J., Qi, J., and Gu, Z. (2021). Characteristics and prognostic value of potential dependency genes in clear cell renal cell carcinoma based on a large-scale CRISPR-Cas9 and RNAi screening database DepMap. *Int. J. Med. Sci.* 18, 2063-2075.

Yang, D., Khan, S., Sun, Y., Hess, K., Shmulevich, I., Sood, A.K., and Zhang, W. (2011). Association of BRCA1 and BRCA2 mutations with survival, chemotherapy sensitivity, and gene mutator phenotype in patients with

ovarian cancer. *JAMA* 306, 1557-1565.

Yang, W., Soares, J., Greninger, P., Edelman, E.J., Lightfoot, H., Forbes, S., Bindal, N., Beare, D., Smith, J.A., Thompson, I.R., et al. (2013). Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 41(Database issue), D955-D961.

Zhong, Z., Hong, M., Chen, X., Xi, Y., Xu, Y., Kong, D., Deng, J., Li, Y., Hu, R., Sun, C., et al. (2020). Transcriptome analysis reveals the link between lncRNA-mRNA co-expression network and tumor immune microenvironment and overall survival in head and neck squamous cell carcinoma. *BMC Med. Genomics* 13, 57.