



CDRgator: An Integrative Navigator of Cancer Drug Resistance Gene Signatures

Su-Kyeong Jang^{1,4}, Byung-Ha Yoon^{2,3,4}, Seung Min Kang^{1,4}, Yeo-Gha Yoon¹, Seon-Young Kim^{2,3,*}, and Wankyu Kim^{1,*}

¹Ewha Research Center for Systems Biology, Department of Life Science, Division of Molecular & Life Sciences, Ewha Womans University, Seoul 03760, Korea, ²Gene Editing Research Center, KRIBB, Daejeon 34141, Korea, ³Department of Functional Genomics, University of Science and Technology (UST), Daejeon 34113, Korea, ⁴These authors contributed equally to this work.
*Correspondence: kimsy@kribb.re.kr (SYK); wkim@ewha.ac.kr (WK)
<http://dx.doi.org/10.14348/molcells.2018.0413>
www.molcells.org

Understanding the mechanisms of cancer drug resistance is a critical challenge in cancer therapy. For many cancer drugs, various resistance mechanisms have been identified such as target alteration, alternative signaling pathways, epithelial-mesenchymal transition, and epigenetic modulation. Resistance may arise via multiple mechanisms even for a single drug, making it necessary to investigate multiple independent models for comprehensive understanding and therapeutic application. In particular, we hypothesize that different resistance processes result in distinct gene expression changes. Here, we present a web-based database, CDRgator (Cancer Drug Resistance navigator) for comparative analysis of gene expression signatures of cancer drug resistance. Resistance signatures were extracted from two different types of datasets. First, resistance signatures were extracted from transcriptomic profiles of cancer cells or patient samples and their resistance-induced counterparts for >30 cancer drugs. Second, drug resistance group signatures were also extracted from two large-scale drug sensitivity datasets representing ~1,000 cancer cell lines. All the datasets are available for download, and are conveniently accessible based on drug class and cancer type, along with analytic features such as clustering analysis, multi-dimensional scaling, and pathway analysis. CDRgator allows meta-analysis of independent resistance models for more comprehensive understanding of drug-resistance mechanisms that is difficult to accomplish with individual datasets alone (database URL: <http://cdrgator.ewha.ac.kr>).

Keywords: cancer drug resistance, gene expression signatures, meta-analysis, microarray, RNA-seq analysis, transcriptome

INTRODUCTION

Overcoming drug resistance is one of the most critical challenges in cancer therapy. Cancer drug resistance appears not only in patients who have genetic factors interfering with drug actions, but is also induced by long-term treatment even in initially responsive patients. Because such acquired resistance leads to significant survival disadvantages, it is important to understand the underlying mechanisms of drug resistance in different patients.

In the early studies on resistance mechanisms, genetic mutations that directly alter drug target proteins (i.e. T790M in EGFR and L1152R in ALK) received significant attention (Choi et al., 2010; Pao et al., 2005). These mutations at gatekeeper residues are a common mechanism of resistance to drugs targeting oncogenic kinases, and many patients develop resistance within a year (Holohan et al., 2013). Recently, mutations other than those in the target protein have also been actively studied including alteration of regulatory regions (Leucci et al., 2018; Melton et al., 2017). Although a number of studies on drug resistance have focused on mutations, it has become clear that various genetic events prior to

Received 21 October, 2018; revised 26 December, 2018; accepted 29 January, 2019; published online 12 February, 2019

eISSN: 0219-1032

© The Korean Society for Molecular and Cellular Biology. All rights reserved.

© 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/>.

the acquisition of resistance also play an important role (Leucci et al., 2018).

In addition to mutations in the target, other genetic or epigenetic changes including copy number variation (CNV, Bean et al., 2007), activation of bypass signaling (Niederst and Engelman, 2013), and epigenetic modulations (Hu and Baeg, 2017; Nyce et al., 1993) have been identified as primary causes of resistance. Resistance mechanisms are not limited to alterations in only a few genes, but involve overall cellular changes, e.g. epithelial-mesenchymal-transition (EMT) or transition to cancer stem cells (CSCs), which seem to be important factors (Huang et al., 2015; Phi et al., 2018; Wang et al., 2016b). Because such major changes in cell state and physiology cannot be represented by a few genetic alterations, investigation of genome-scale signatures such as transcriptome profiles is likely to be more informative. Transcriptome profiles reflect changes in both genetic and epigenetic conditions including disease progression, drug perturbation, and drug resistance. Therefore, gene expression signatures are useful markers to investigate drug resistance states, classify distinct resistance mechanisms, and predict treatment outcomes.

Typically, there are two strategies to study genetic features associated with drug sensitivity. One is to use large-scale pharmacogenomic datasets of cancer cell lines as model systems, with the cells being divided into two groups, i.e. sensitive and resistant, and genomic features differentiating the two groups being identified. The other is to culture cancer cells in the presence of a drug, leading to selection of resistant cells. By comparing the molecular profiles of the resistance-induced cells with those of the original ones, the relevant genomic features can be analyzed more explicitly than by using the group-based method because the origins of the cells are the same in the former method.

There are several publicly available pharmacogenomics databases such as Genomics of Drug Sensitivity in Cancer (GDSC, <https://www.cancerrxgene.org/>) (Yang et al., 2013), Cancer Cell Line Encyclopedia (CCLE, <https://portals.broadinstitute.org/ccle/>) (Barretina et al., 2012), and Cancer Therapeutic Response Portal (CTRP, extended version of CCLE, <https://portals.broadinstitute.org/ctrp.v2.1/>) (Rees et al., 2016). From studies of these databases, mutations in or expression of individual genes associated with drug sensitivity have been characterized. These features have been found useful in predicting resistance based on the genetic background of patients (Iorio et al., 2016; Seashore-Ludlow et al., 2015). While these datasets are among the largest resources to study drug resistance mechanisms, the sparsity of individual mutations makes it difficult to extract genetic markers related to resistance in a comprehensive manner. Only a small fraction of all acquired resistance events in patients may be explained using the markers extracted from GDSC, CCLE, CTRP, and the combination dataset. For example, though the T790M mutation in the EGFR gene has been well known as one of the most frequent alterations associated with acquired resistance to EGFR-targeted drugs in patients (Ma et al., 2011), this mutation has been found in only two cancer cell lines in the above pharmacogenomics databases. Although the Catalog of Somatic Mutations In Cancers

(COSMIC, <https://cancer.sanger.ac.uk/cosmic/>) (Forbes et al., 2017) has some genetic features associated with acquired resistance, there is no systematic database of information regarding acquired resistance yet. In COSMIC, mutation information is available for only 13 genes associated with resistance to 23 drugs. There are other databases unrelated to cancer therapy such as databases of antibiotic or antimicrobial drug resistance such as CARD (Jia et al., 2017) and MEGARes (Lakin et al., 2017), respectively. Notably, cancerDR (Kumar et al., 2013) may be the only database dedicated to cancer drug resistance. However, because the resistance-related features in cancerDR mostly originate from CCLE and COSMIC, it is likely to have limitations similar to those of its original resources.

In the present study, we present a web platform, CDRgator (Cancer Drug Resistance signature navigator), which is among the most comprehensive databases of acquired cancer drug resistance because it amasses both *resistance-induced signatures* (S_{ind}) and *group-based signatures* (S_{grp}) of drug sensitivity, with the S_{ind} defined as the differential gene expression in resistant cells of the same lineage induced due to prolonged culture of a cell line in the presence of drug, and S_{grp} is defined as that in resistant or sensitive groups of cells from different origins. Due to the limitations in genomic features due to sparse mutations, we focused on extracting gene expression signatures of cancer drug resistance to obtain a comprehensive view of resistance mechanisms. For constructing resistance signatures of resistance-induced cells (S_{ind}), we performed extensive manual curation of literature as well as data depositories such as Gene express omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) (Barrett et al., 2013) and ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>) (Kolesnikov et al., 2015). Alternatively, we also extracted group-based resistance signatures for cancer drugs from CCLE and CTRP; we refer to these resistance signatures extracted from the two large-scale pharmacogenomics databases as *group-based signatures* (S_{grp}). These two types of signatures are complementary to each other, and are expected to provide a more comprehensive view on drug resistance, particularly when there are multiple, independent mechanisms involved in developing resistance to a single drug.

Currently, CDRgator provides 603 resistance signatures for 37 cancer drugs representing more than 26 cancer types in total, and the number of signatures will grow as more data are collected. It allows users to browse resistance signatures based on drugs or cancer types, and to analyze the similarity between resistance signatures. Additionally, CDRgator has a tool to identify resistance signature-matched gene sets from Kyoto Encyclopedia of Genes and Genomes (KEGG, (Kanehisa, 2004) or Gene ontology (GO) (Carbon et al., 2017) to characterize the biological processes involved in resistance. CDRgator also has the ability to compare the resistance signatures in a database with signatures input by users to filter out drugs expected to be resistant or ineffective. Using an illustrative analysis of EGFR inhibitors, we show the utility of CDRgator in understanding the diverse mechanisms of cancer drug resistance.

MATERIALS AND METHODS

Data sources and processing

To obtain induced drug resistance signatures (S_{ind}), we manually collected datasets containing transcriptomic profiles of resistance-induced cells in GEO and ArrayExpress (Fig. 1). The datasets were filtered based on the following criteria; (1) the presence of matched pairs of sensitive and resistant cells, (2) treatment with monotherapy but not combinational therapy for a specific period of time, (3) gene transcription quantification using RNA-seq or microarray (4) in human cells and not in mouse cells or xenografted mouse cells. The RNA-seq data (.fastq file) were mapped using STAR aligner (Dobin et al., 2013) and quantified using HTSeq (Anders et al., 2015). The list of the collected datasets is available with detailed information in Additional file 1.

For identifying the resistant group signature (S_{grp}), we used drug sensitivity data and gene expression data of cancer cell lines from CTRP and CCLE, respectively. To designate the resistant cell line group for a given drug, we grouped cancer

cell lines based on their drug responses. The drug responses were represented as area under curve (AUC) of cell growth at different drug concentrations, and normalized with the maximum area calculated assuming 100% response in given concentration ranges. Then, we performed Z-transformation of logged AUC values (Z-AUC) to specify resistant or sensitive cells. We defined cells as a resistant cell group if their Z-AUC was less than -0.5. For all drugs, we generated signatures only in cases with at least three cells in the resistant group.

The expression data were downloaded from GEO or ArrayExpress, and were processed using our protocols for microarray and RNAseq analyses. Background correction and normalization was performed using R. For RNAseq, raw fastq files were downloaded and aligned to the reference genome (GRCh37) using STAR aligner. For RNAseq of cancer cell lines, BAM files were downloaded from Genomic Data Commons (GDC). To estimate expression, RNAseq read count was calculated using HTSeq. The gene expression signature was obtained using differentially expressed gene

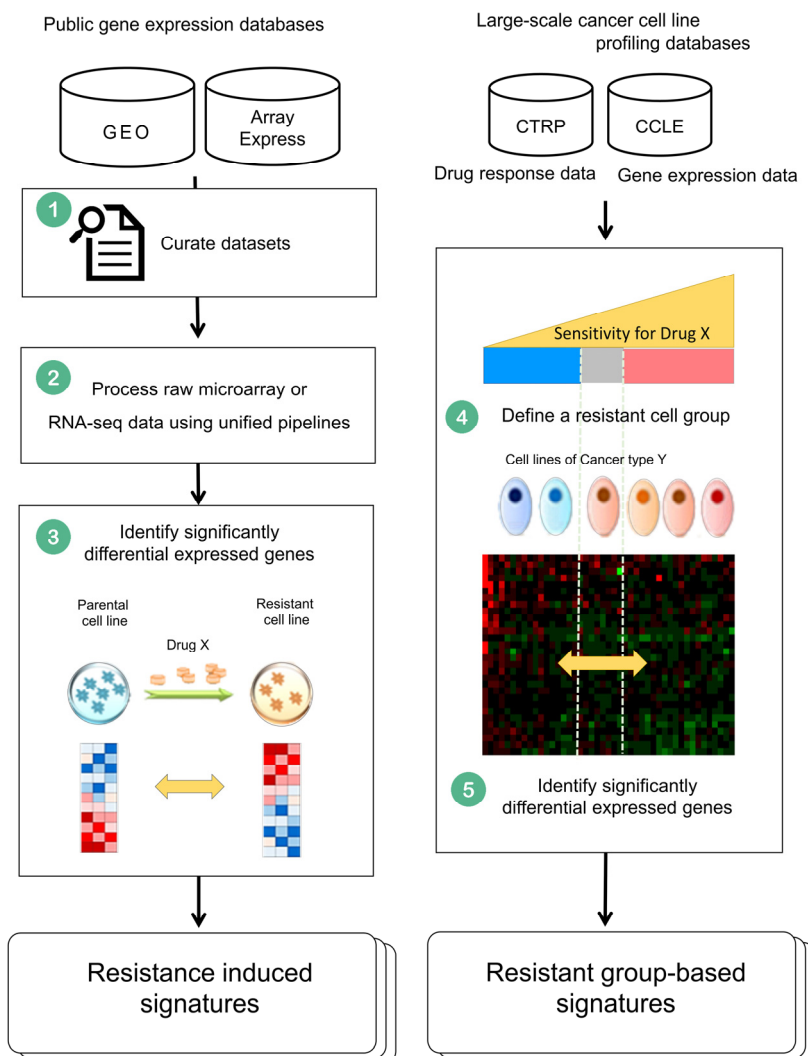


Fig. 1. The process of extracting drug resistance signatures. Left, resistance-induced signatures; Right, resistance group signatures.

(DEG) analysis of the resistant cell group or the resistance-induced cells. For datasets containing biological replicates, genes with adjusted p-values less than 0.05 were selected as DEGs using the limma package (Ritchie et al., 2015) for microarray, and the DESeq2 package (Love et al., 2014) for RNAseq. For datasets with no biological replicates, we calculated fold change in gene expression values and transformed them to z-score. Genes with absolute fold change in z-score of more than 2 were selected as DEGs.

Data statistics

CDRgator contains a total of 143 S_{ind} for 37 anti-cancer drugs and 30 cell lines from 16 tissue origins. A set of 499 S_{grp} in our study represented 19 drugs which were also included in S_{ind} ; in terms of details, 267 S_{grp} were from 20 tissue-level cancer types (e.g. blood cancer) and 232 S_{grp} were from 23 disease level-cancer types (e.g. acute lymphoblastic leukemia).

Signature similarity

CDRgator provides similarity analysis which explores how similar two signatures are; an enrichment factor (EF) was

developed to measure significance of overlap between the signatures compared to a random occurrence. It represents the number of genes common between two sets divided by the expected overlap. At this time, every gene set is restricted to gene spaces common to respective platforms of a signature pair. EF is calculated using the following equation:

$$\text{Enrichment factor (A, B)} = \frac{|A \cap B| + pc}{N \times \frac{|A|}{N} \times \frac{|B|}{N} + pc} = \frac{|A \cap B| + pc}{\frac{|A| \times |B|}{N} + pc}$$

where A and B indicate a set of signature genes, and N represents the number of total genes common to both platforms used; pc is a pseudo count that makes the EF robust at small count levels (5 in this study), similar to additive smoothing. EF has the particular advantage of being robust despite differences in signature size.

Functional annotation analysis

To annotate the biological functions of gene signatures, an over-representation test was conducted using the 'clusterProfiler' in the R package (Yu et al., 2012), which implements a hypergeometric test of overrepresentation of a gene

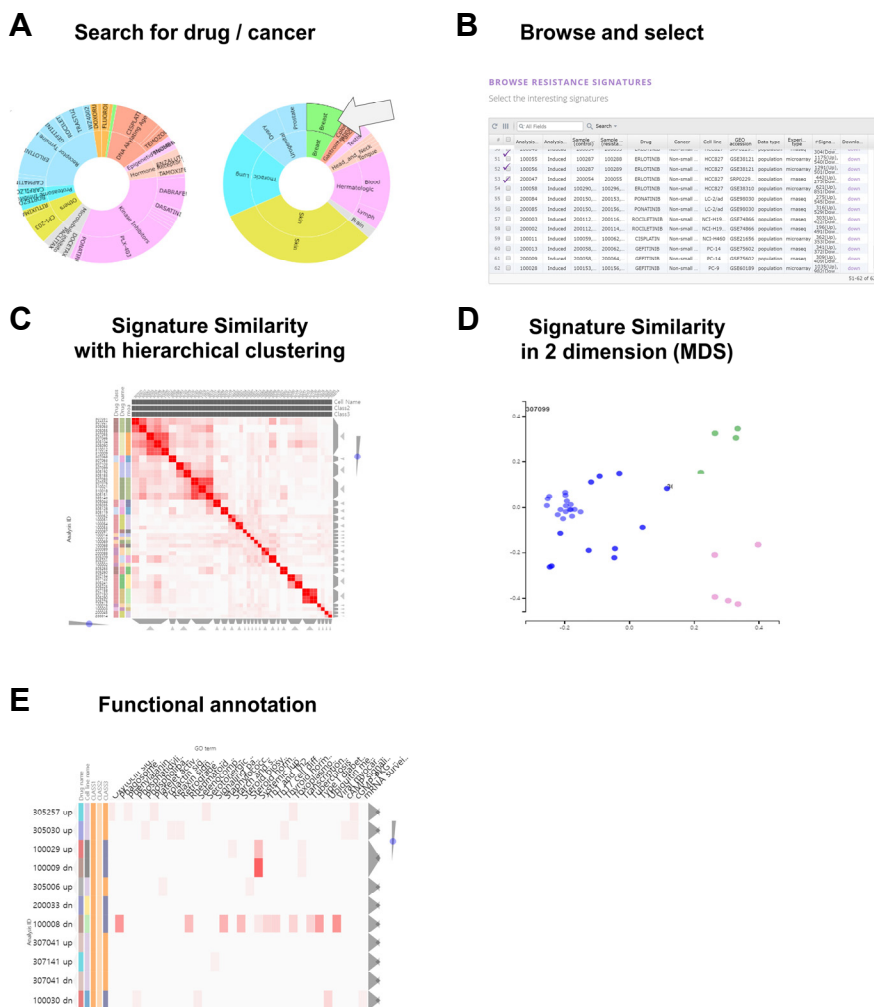


Fig. 2. Web interface design of CDRgator. (A) selection of drug class (left circle) and cancer type (right circle); (B) selection of individual datasets; (C) hierarchical clustering of resistance signatures; (D) 2D mapping of resistance signatures via multidimensional scaling (MDS); (E) pathway analysis of resistance signatures using gene set analysis (GSA).

set against background. Each signature analysis was performed with pathway terms consisting of 10-500 genes among the KEGG pathways or GO biological processes. As the size of the signature set decreases, the number of significantly enriched pathways decreases. Thus, CDRgator also generates a p-value representing the significance of functional enrichment.

User interface

The CDRgator web service is freely available at <http://cdrgator.ewha.ac.kr>. Extraction of signatures from expression data and analysis of CDRgator web platform are implemented in R script (R version 3.4.4). All the generated data were stored in the MySQL database (v5.7.22). Control of the data from a database has been implemented in Java (JRE 1.8.0) and we built a user-friendly web interface using JSP and jQuery to conveniently provide the data to the user. In addition, the CDRgator web interface uses html5, css3, SASS, and jQuery to load content intuitively and quickly on a variety of platforms, including desktop and mobile. For rich and intuitive visualization of data, it uses the clustergrammer (Fernandez et al., 2017) which is based on the D3.js JavaScript library to

build an advanced web interface with animation. The help documentation provides a video and a brief description of the CDRgator guide.

CDRgator is primarily composed of 'Browse' and 'Analysis' menus. The 'Browse' menu provides the ability to search and browse information regarding resistance signatures. The 'Analysis' menu performs signature similarity and functional analysis. Signature similarity analysis calculates how similar resistance signatures are, so that users can identify if signatures of specific cancers, drugs, or their classes show a significant tendency in original spaces or two dimensions through multidimensional scaling. A functional analysis shows a heatmap that is generated from an over-representation test to identify genetic and biological functions of signatures. Finally, significant genes can be visualized using KEGG pathways (Supplementary Fig. S1).

RESULTS

Analysis of EGFR inhibitor resistance-induced signatures

To demonstrate the utility of CDRgator, we performed a case study of induced EGFR inhibitor resistance signatures.

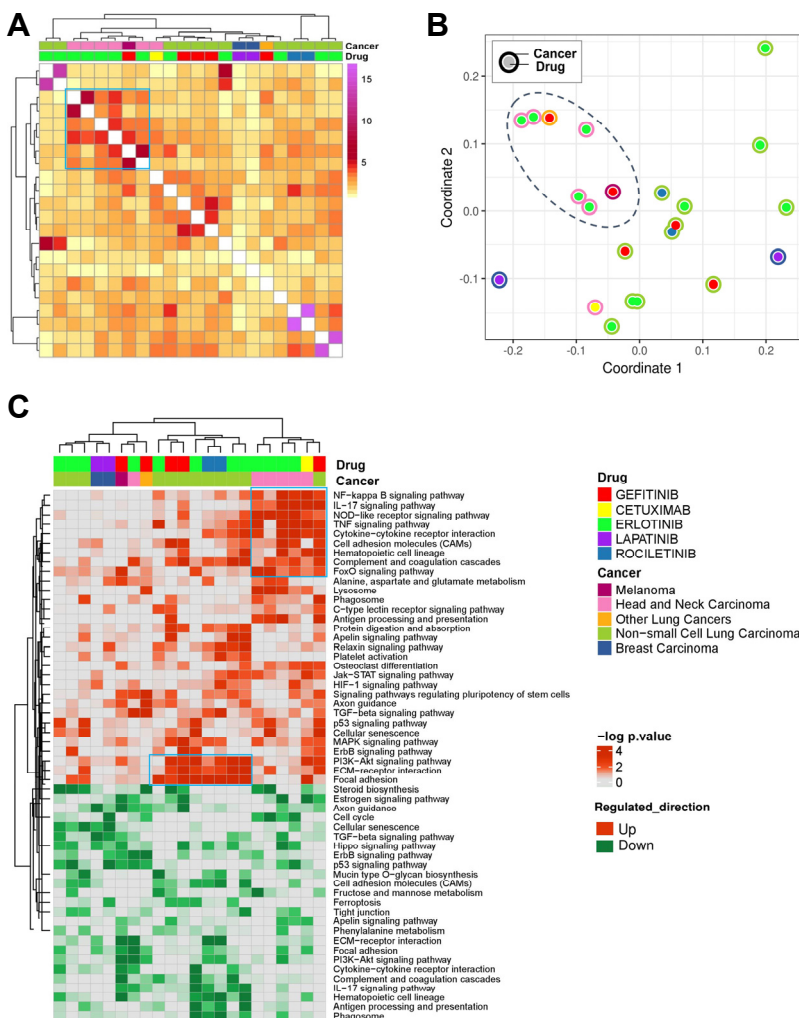


Fig. 3. Analysis of resistance-induced signatures for EGFR inhibitors. Illustrative case results available in the 'Analysis' menu of CDRgator, generated using resistance-induced signatures for five EGFR inhibitors. Color legends of drug and cancer type are described in Fig. 3C; (A) signature similarity analysis; heat represents all pairwise similarities; hierarchical clustering of similarity is shown at the top and left; (B) multidimensional scaling plot of similarity metrics; outer circles are colored cancer types, and inner circles are drugs; the dashed line indicates a cluster including primarily head and neck cancer signatures; (C) KEGG pathway enrichment analysis; up (red) or down (green) regulated gene sets in individual signatures were tested using each pathway term on the left; heat indicates $-\log p$ -value of hypergeometric testing between each pathway gene set and signatures; hierarchical clustering is shown on the top and left.

There were 22 induced resistance signatures for five different EGFR inhibitors. Because these signatures were generated independent of each other, the following questions could be asked; i) is the resistance pattern dependent on the drug, the cancer type, or both; and ii) do drugs with similar mode-of-action (e.g. EGFR inhibitors) share changes in transcriptomic profiles.

Similarity clustering analysis of the 22 S_{ind} for EGFR inhibitors showed a tight cluster, with mostly erlotinib resistance signatures from head and neck cancer (Fig. 3A). This suggests distinct and reproducible patterns of resistance mechanisms specific to a drug and cancer type, because similar signature patterns were generated from multiple independent experiments. This trend was also confirmed in 2D mapping of the signatures using multidimensional scaling (Fig. 3B). These signatures were distinguishable from those of other cancer types. Overall, the resistance patterns collected in CDRgator appear to be heterogeneous even for the same drug and cancer type (e.g. erlotinib (center green) in non-small cell lung carcinoma (NSCLC, green outer ring) in Fig. 3B).

Pathway analysis showed the heterogeneity in resistance signatures in more detail. In head and neck cancers, biological pathways previously known to be associated with drug resistance were also enriched in our signatures, such as up-regulation of NF- κ B, NOD-like receptor, TNF, and FoxO signaling pathways. NF- κ B signaling is well-known among the survival and resistance pathways against many anticancer drugs including EGFR-targeted drugs (Bentires-Alj et al., 2003; Hertlein and Byrd, 2010; Lagunas and Meléndez-Zajgla, 2008; Wang et al., 2018). It is known that the NF- κ B activating complex drives cancer cells to become resistant against EGFR inhibitors (Blakely et al., 2015), and this signaling pathway could also lead to resistance against EGFR inhibitors designed to target mutated forms of EGFR (Galvani et al., 2015). In colon cancers, activated NF- κ B induces drug resistance through the regulation of MDR1 gene expression (Bentires-Alj et al., 2003). EGFR triggers proto-oncogenic signals such as RAS and mitogen activated protein kinase (MAPK) signaling (Wee and Wang, 2017). NOD-like receptors are known to regulate innate immunity by forming inflammasomes, and also by activating MAPK and NF- κ B signaling (Saxena and Yeretssian, 2014; Shaw et al., 2010). NOD1/2 downstream signaling was reported to confer drug resistance through RIP2-PAX5 interaction by activating MAPK and NF- κ B signaling (Wang et al., 2016a). AKT-PI3K signaling is also known to be activated in EGFR drug-resistant cancer cells (Jacobsen et al., 2017; Ma et al., 2016). FOXO is among the downstream genes playing a role in cell survival and apoptosis. Drug sensitivity was restored via inhibition of FOXO in AKT-mediated EGFR inhibitor-resistant lung cancers (Sangodkar et al., 2012). Several of these pathways (i.e. NF- κ B, NOD-like receptor, and FOXO signaling) were also frequently upregulated in other resistance signatures (Fig. 3B).

In NSCLC, additional pathways (i.e. PI3K-Akt, ECM-receptor interaction, and focal adhesion) were strongly enriched. These pathways were found to be associated with epithelial-mesenchymal transition (EMT), which has been

recently highlighted as major resistance mechanism in both chemotherapy and targeted cancer drug therapy (Chen et al., 2013; Huang et al., 2015; Larue and Bellacosa, 2005; Le Bras et al., 2012; Wang et al., 2016b). As shown in the example with EGFR inhibitors, CDRgator provides rich information on the drug resistance processes in terms of heterogeneity, and dependence on cancer type and drug class.

DISCUSSION

The development of technology coupled with bioinformatics and system biology approaches has enabled the identification of genomic and transcriptome features to predict response and resistance to specific drugs. The wide range of molecular mechanisms involving these features has important implications for understanding and treating resistance. However, inhibiting one pathway can result in a relatively simple escape route for the tumor, and an integrated analysis approach involving various features to overcome or circumvent drug resistance is needed.

CDRgator was developed for studying the acquired resistance mechanisms against cancer drugs. Gene expression signatures were generated from manually collected data regarding cancer drug resistance-induced cells in public datasets. CDRgator allows inspection and comparison of these acquired resistance transcriptomic signatures. Pathway analysis enables functional interpretation of multiple signatures simultaneously. While the datasets available in CDRgator are still sparse and insufficient to provide comprehensive information on drug resistance mechanisms, it may provide a useful clue to investigate drug resistance mechanisms, and to develop the right therapeutic strategy. For example, an analysis of genome-scale transcript signatures provided by CDRgator and recurrent mutations in the regulatory region may provide a variety of data enabling resistance marker identification or mechanistic analysis. The complete datasets in CDRgator are accessible at <http://cdrgator.ewha.ac.kr> and are also downloadable as mysql database dump.

Note: Supplementary information is available on the Molecules and Cells website (www.molcells.org).

ACKNOWLEDGMENTS

This work was funded by grants from National Research Foundation of Korea (NRF-2017R1A2B4007855, NRF-2017M3C9A5028690, and NRF-2014M3C9A3068554) and from Institution for Information & communications Technology Promotion (IITP) (No. 2016-0-00289).

REFERENCES

- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq-A python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166-169.
- Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, 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.
- Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F.,

- Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., et al. (2013). NCBI GEO: archive for functional genomics data sets - update. *Nucleic Acids Res.* *41*, 991-995.
- Bean, J., Brennan, C., Shih, J.-Y., Riely, G., Viale, A., Wang, L., Chitale, D., Motoi, N., Szoke, J., Broderick, S., et al. (2007). MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc. Natl. Acad. Sci.* *104*, 20932-20937.
- Bentires-Alj, M., Barbu, V., Fillet, M., Chariot, A., Relic, B., Jacobs, N., Gielen, J., Merville, M.P., and Bours, V. (2003). NF- κ B transcription factor induces drug resistance through MDR1 expression in cancer cells. *Oncogene* *22*, 90-97.
- Blakely, C.M., Pazarentzos, E., Olivas, V., Asthana, S., Yan, J.J., Tan, I., Hrustanovic, G., Chan, E., Lin, L., Neel, D.S., et al. (2015). NF- κ B-Activating complex engaged in response to EGFR oncogene inhibition drives tumor cell survival and residual disease in lung cancer. *Cell Rep.* *11*, 98-110.
- Le Bras, G.F., Taubenslag, K.J., and Andl, C.D. (2012). The regulation of cell-cell adhesion during epithelial-mesenchymal transition, motility and tumor progression. *Cell Adhes. Migr.* *6*, 365-373.
- Carbon, S., Dietze, H., Lewis, S.E., Mungall, C.J., Munoz-Torres, M.C., Basu, S., Chisholm, R.L., Dodson, R.J., Fey, P., Thomas, P.D., et al. (2017). Expansion of the gene ontology knowledgebase and resources: the gene ontology consortium. *Nucleic Acids Res.* *45*, D331-D338.
- Chen, Q.K., Lee, K., Radisky, D.C., and Nelson, C.M. (2013). Extracellular matrix proteins regulate epithelial-mesenchymal transition in mammary epithelial cells. *Differentiation* *86*, 126-132.
- Choi, Y.L., Soda, M., Yamashita, Y., Ueno, T., Takashima, J., Nakajima, T., Yatabe, Y., Takeuchi, K., Hamada, T., Haruta, H., et al. (2010). EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors. *N. Engl. J. Med.* *363*, 1734-1739.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* *29*, 15-21.
- Fernandez, N.F., Gundersen, G.W., Rahman, A., Grimes, M.L., Rikova, K., Hornbeck, P., and Maayan, A. (2017). Clustergrammer, a web-based heatmap visualization and analysis tool for high-dimensional biological data. *Sci. Data* *4*, 1-12.
- Forbes, S.A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C.G., Ward, S., Dawson, E., Ponting, L., et al. (2017). COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res.* *45*, D777-D783.
- Galvani, E., Sun, J., Leon, L.G., Sciarriello, R., Narayan, R.S., Sjin, R.T.T., Lee, K., Ohashi, K., Heideman, D.A.M., Alfieri, R.R., et al. (2015). NF- κ B drives acquired resistance to a novel mutant-selective EGFR inhibitor. *Oncotarget* *6*, 42717-42732.
- Hertlein, E., and Byrd, J.C. (2010). Signalling to drug resistance in CLL. *Best Pract. Res. Clin. Haematol.* *23*, 121-131.
- Holohan, C., Schaeuble, S.V., Longley, D.B., and Johnston, P.G. (2013). Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer* *13*, 714.
- Hu, Q., and Baeg, G.H. (2017). Role of epigenome in tumorigenesis and drug resistance. *Food Chem. Toxicol.* *109*, 663-668.
- Huang, J., Li, H., and Ren, G. (2015). Epithelial-mesenchymal transition and drug resistance in breast cancer (Review). *Int. J. Oncol.* *47*, 840-848.
- Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., et al. (2016). A landscape of pharmacogenomic interactions in cancer. *Cell* *166*, 740-754.
- Jacobsen, K., Bertran-Alamillo, J., Molina, M.A., Teixidó, C., Karachaliou, N., Pedersen, M.H., Castellví, J., Garzón, M., Codony-Servat, C., Codony-Servat, J., et al. (2017). Convergent Akt activation drives acquired EGFR inhibitor resistance in lung cancer. *Nat. Commun.* *8*, 410.
- Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., Lago, B.A., Dave, B.M., Pereira, S., Sharma, A.N., et al. (2017). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* *45*, D566-D573.
- Kanehisa, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res.* *32*, 277D-280.
- Kolesnikov, N., Hastings, E., Keays, M., Melnichuk, O., Tang, Y.A., Williams, E., Dylag, M., Kurbatova, N., Brandizi, M., Burdett, T., et al. (2015). ArrayExpress update-simplifying data submissions. *Nucleic Acids Res.* *43*, D1113-D1116.
- Kumar, R., Chaudhary, K., Gupta, S., Singh, H., Kumar, S., Gautam, A., Kapoor, P., and Raghava, G.P.S. (2013). CancerDR: cancer drug resistance database. *Sci. Rep.* *3*, 1445.
- Lagunas, V.M., and Meléndez-Zajgla, J. (2008). Nuclear factor-kappa B as a resistance factor to platinum-based antineoplastic drugs. *Met. Based. Drugs* *2008*, 576104.
- Lakin, S.M., Dean, C., Noyes, N.R., Dettenwanger, A., Ross, A.S., Doster, E., Rovira, P., Abdo, Z., Jones, K.L., Ruiz, J., et al. (2017). MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res.* *45*, D574-D580.
- Larue, L., and Bellacosa, A. (2005). Epithelial-mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways. *Oncogene* *24*, 7443-7454.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* *15*, 1-21.
- Leucci, E. (2018). Cancer development and therapy resistance: spotlights on the dark side of the genome. *Pharmacol Ther.* *189*, 22-30.
- Ma, C., Wei, S., and Song, Y. (2011). T790M and acquired resistance of EGFR TKI: a literature review of clinical reports. *J. Thorac. Dis.* *3*, 10-18.
- Ma, P., Fu, Y., Chen, M., Jing, Y., Wu, J., Li, K., Shen, Y., Gao, J.X., Wang, M., Zhao, X., et al. (2016). Adaptive and acquired resistance to EGFR inhibitors converge on the MAPK pathway. *Theranostics* *6*, 1232-1243.
- Melton, C., Reuter, J.A., Spacek, D.V., and Snyder, M. (2017). Recurrent somatic mutations in regulatory regions of human cancer genomes. *Nature Genetics* *47*, 710.
- Niederst, M.J., and Engelman, J.A. (2013). Bypass mechanisms of resistance to receptor tyrosine kinase inhibition in lung cancer. *Sci. Signal.* *6*, re6-re6.
- Nyce, J., Leonard, S., Canupp, D., Schulz, S., and Wong, S. (1993). Epigenetic mechanisms of drug resistance: drug-induced DNA hypermethylation and drug resistance. *Proc. Natl. Acad. Sci. USA* *90*, 2960-2964.
- Pao, W., Miller, V.A., Politi, K.A., Riely, G.J., Somwar, R., Zakowski, M.F., Kris, M.G., and Varmus, H. (2005). Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med.* *2*, e7.
- Phi, L.T.H., Sari, I.N., Yang, Y.-G., Lee, S.-H., Jun, N., Kim, K.S., Lee, Y.K., and Kwon, H.Y. (2018). Cancer stem cells (CSCs) in drug resistance and their therapeutic implications in cancer treatment. *Stem Cells Int.* *2018*, 1-16.
- Rees, M.G., Seashore-Ludlow, B., Cheah, J.H., Adams, D.J., Price, E.

- V., Gill, S., Javaid, S., Coletti, M.E., Jones, V.L., Bodycombe, N.E., et al. (2016). Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat. Chem. Biol.* *12*, 109-116.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* *43*, e47.
- Sangodkar, J., Difeo, A., Narla, G., Sangodkar, J., Dhawan, N.S., Melville, H., Singh, V.J., Yuan, E., Rana, H., Izadmehr, S., et al. (2012). Targeting the FOXO1 / KLF6 axis regulates EGFR signaling and treatment response Find the latest version: targeting the FOXO1 / KLF6 axis regulates EGFR signaling and treatment response. *122*, 2637-2651.
- Saxena, M., and Yeretssian, G. (2014). NOD-like receptors: master regulators of inflammation and cancer. *Front. Immunol.* *5*, 1-16.
- Seashore-Ludlow, B., Rees, M.G., Cheah, J.H., Cokol, M., Price, E. V., Coletti, M.E., Jones, V., Bodycombe, N.E., Soule, C.K., Gould, J., et al. (2015). Harnessing connectivity in a large-scale small-molecule sensitivity dataset. *Cancer Discov.* *5*, 1210-1223.
- Shaw, P.J., Lamkanfi, M., and Kanneganti, T.D. (2010). NOD-like receptor (NLR) signaling beyond the inflammasome. *Eur. J. Immunol.* *40*, 624-627.
- Wang, D., Chen, J., Li, R., Wu, G., Sun, Z., Wang, Z., Zhai, Z., Fang, F., Guo, Y., Zhong, Y., et al. (2016a). PAX5 interacts with RIP2 to promote NF- κ B activation and drug-resistance in B-lymphoproliferative disorders. *J. Cell Sci.* *129*, 2261-2272.
- Wang, J., Wei, Q., Wang, X., Tang, S., Liu, H., Zhang, F., Mohammed, M.K., Huang, J., Guo, D., Lu, M., et al. (2016b). Transition to resistance: an unexpected role of the EMT in cancer chemoresistance. *Genes Dis.* *3*, 3-6.
- Wang, X., Yin, H., Zhang, H., Hu, J., Lu, H., Li, C., Cao, M., Yan, S., and Cai, L. (2018). NF- κ B-driven improvement of EHD1 contributes to erlotinib resistance in EGFR-mutant lung cancers. *Cell Death Dis.* *9*, 418.
- Wee, P., and Wang, Z. (2017). Epidermal growth factor receptor cell proliferation signaling pathways. *Cancers (Basel)*. *9*, 1-45.
- 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*, 955-961.
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* *16*, 284-287.