Omics and Computational Modeling Approaches for the Effective Treatment of Drug-Resistant Cancer Cells

Hae Deok Jung¹, Yoo Jin Sung¹ and Hyun Uk Kim¹,2,3*

¹Department of Chemical and Biomolecular Engineering (BK21 four), Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, ²KAIST Institute for Artificial Intelligence, KAIST, Daejeon, South Korea, ³BioProcess Engineering Research Center and BioInformatics Research Center KAIST, Daejeon, South Korea

Chemotherapy is a mainstream cancer treatment, but has a constant challenge of drug resistance, which consequently leads to poor prognosis in cancer treatment. For better understanding and effective treatment of drug-resistant cancer cells, omics approaches have been widely conducted in various forms. A notable use of omics data beyond routine data mining is to use them for computational modeling that allows generating useful predictions, such as drug responses and prognostic biomarkers. In particular, an increasing volume of omics data has facilitated the development of machine learning models. In this mini review, we highlight recent studies on the use of multi-omics data for studying drug-resistant cancer cells. We put a particular focus on studies that use computational models to characterize drug-resistant cancer cells, and to predict biomarkers and/or drug responses. Computational models covered in this mini review include network-based models, machine learning models and genome-scale metabolic models. We also provide perspectives on future research opportunities for combating drug-resistant cancer cells.

Keywords: cancer, drug resistance, omics, computational modeling, network-based model, machine learning, genome-scale metabolic model

INTRODUCTION

Drug resistance has been a major obstacle for a successful treatment of cancers, as manifested by over 90% mortality of cancer patients that appeared to be associated with drug resistance (Bukowski et al., 2020). Drug resistance is a phenotypic state that arises as a result of a complex interplay between genetic and non-genetic mechanisms (Marine et al., 2020). Such genetic and non-genetic reprogramming consequently leads to drug resistance through various mechanisms (Gatti and Zunino, 2005; Housman et al., 2014; Zheng, 2017; Lim and Ma, 2019; Vasan et al., 2019; Bukowski et al., 2020), including: drug inactivation, for example by an excessive level of glutathione that detoxifies xenobiotics (Jiang et al., 2017; De Luca et al., 2019); alteration of a drug target by mutations or changes in an expression level (Likhite et al., 2006; Costa et al., 2008); drug efflux by transporters (Giddings et al., 2021); enhanced DNA damage repair system (Harte et al., 2014); development of resistance via dysregulated autophagy (Martin et al., 2017; Cai et al., 2019); epithelial-mesenchymal transition (EMT) (Fischer et al., 2015; Zheng et al., 2015); or heterogeneity of a cancer cell population having cancer...
stem cells (Seth et al., 2019; Zhao et al., 2021). A state of drug resistance is indeed a highly complex phenotype that requires multidimensional approaches.

Omics technologies have now become indispensable for characterizing mechanisms of cancer progression, and for identifying effective biomarkers and treatment targets for cancers. For this reason, large-scale projects have been launched to generate omics data of various cancer cells. A recent representative example is the Pan-Cancer Analysis of Whole Genomes (PCAWG) Consortium of the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA), which has allowed advanced studies on gene mutations and gene expression profiles across cancers (Consortium, 2020). The resulting various datasets from such large-scale efforts have been found to be useful for studying drug-resistant cancer cells. Relevant representative datasets include the NCI-60 Human Tumor Cell Lines Screen (Shoemaker, 2006), the Genomics of Drug Sensitivity in Cancer (GDSC) (Yang et al., 2013), TCGA (Cancer Genome Atlas Research et al., 2013), the Cancer Therapeutic Response Portal (CTRP) (Seashore-Ludlow et al., 2015), L1000 profiles from The Library of Integrated Network-Based Cellular Signatures (LINCS) Program (Subramanian et al., 2017), the Cancer Cell Line Encyclopedia (CCLE) (Ghandi et al., 2019), and the Catalogue Of Somatic Mutations In Cancer (COSMIC) (Tate et al., 2019). All these datasets have served as a source of novel insights that help characterize and overcome drug-resistant cancer cells. In particular, it is expected that an increasing volume of such large-scale datasets will facilitate development of various computational models that will better systematize our approaches to studying drug-resistant cancer cells.

We here review recent studies that utilized multi-omics and computational modeling approaches to better understand mechanisms associated with the progression of drug resistance, and to identify biomarkers and/or drug responses (Figure 1 and Table 1). Especially, we put more focus on computational modeling that makes predictions for various scenarios for the treatment of drug-resistant cancer cells. We also provide an outlook for further advances on the use of computational models for studying drug-resistant cancer cells.

### MULTI-OMICS ANALYSES

Multiple omics data are often generated to examine various biological aspects of drug-resistant cancer cells (Figure 1). Target genotypes and phenotypes examined using omics data (Table 1) include: cancer-associated mutations (Niehr et al., 2018; Marczyk et al., 2020; Sinkala et al., 2021); changes in the expression level of specific genes (Niehr et al., 2018; Nava et al., 2019; Kagohara et al., 2020; Marczyk et al., 2020; Poojan et al., 2020; Sinkala et al., 2021); changes in chromosome structure (Kagohara et al., 2020; Marczyk et al., 2020; Aissa et al., 2021); epigenetic alterations (e.g., methylation or acetylation states of histone
**TABLE 1 |** Recent studies on the use of omics data and computational models to better understand and treat drug-resistant cancer cells.

| Approaches                              | Cancer types                  | Resistance type                     | Objectives                                      | Drugs                                                                 | References            |
|-----------------------------------------|-------------------------------|-------------------------------------|------------------------------------------------|----------------------------------------------------------------------|-----------------------|
| • ChIP-seq                              | • Lung cancer                 | • Both acquired and intrinsic resistance | • Identification of biomarkers              | Erlotinib, osimertinib, crizotinib, vemurafenib, celestrol, and GSK-1059615 | Aissa et al. (2021)    |
| • Single-cell RNA-seq                   |                               |                                     |                                                |                                                                      |                       |
| • RNA-seq                               |                               |                                     |                                                |                                                                      |                       |
| • Proteome (LC-MS/MS)                   |                               |                                     |                                                |                                                                      |                       |
| • ATAC-seq                              | • Breast cancer               | • Intrinsic resistance              | • Biological characterization                  | Doxorubicin                                                          | Kumar et al. (2021)   |
| • RNA-seq                               |                               |                                     |                                                |                                                                      |                       |
| • Genome sequencing                     | 101 Types of cancers from 40,848 patients from cBioPortal | • Not specified                      | • Biological characterization                  | MAPK pathway inhibitors (e.g., selumetinib)                         | Sinkala et al. (2021) |
| • Methylation (reduced representation bisulfite sequencing) |                               |                                     |                                                |                                                                      |                       |
| • mRNA microarray and RNA-seq           |                               |                                     |                                                |                                                                      |                       |
| • RNA-seq                               | • Melanoma                    | • Intrinsic resistance              | • Biological characterization                  | Verumafenib                                                          | Torre et al. (2021)   |
| • Pooled CRISPR screen (MiSeq)          |                               |                                     |                                                |                                                                      |                       |
| • Genome sequencing                     | • Breast cancer               | • Acquired resistance               | • Biological characterization                  | Tamoxifen and fulvestrant                                            | Achinger-Kawecka et al. (2020) |
| • Methylation (bisulfite sequencing)    |                               |                                     |                                                |                                                                      |                       |
| • Hi-C                                  |                               |                                     |                                                |                                                                      |                       |
| • ChIP-seq                              |                               |                                     |                                                |                                                                      |                       |
| • RNA-seq                               |                               |                                     |                                                |                                                                      |                       |
| • Methylome (EPIC array)                | • Breast cancer               | • Acquired resistance               | • Biological characterization                  | Paclitaxel                                                          | DeBlois et al. (2020) |
| • ChIP-seq                              |                               |                                     |                                                |                                                                      |                       |
| • RNA-seq                               |                               |                                     |                                                |                                                                      |                       |
| • Metabolome (LC-HRMS)                  |                               |                                     |                                                |                                                                      |                       |
| • ATAC-seq                              | • Head and neck squamous carcinoma | • Acquired resistance               | • Biological characterization                  | Gatteximab                                                          | Kagohara et al. (2020) |
| • Single-cell RNA-seq                   |                               |                                     |                                                |                                                                      |                       |
| • RNA-seq                               |                               |                                     |                                                |                                                                      |                       |
| • Translatome (microarray)              | • Leukemia                    | • Not specified                      | • Biological characterization                  | Cytosine arabinoside                                                | Lee et al. (2020)     |
| • mRNA microarray                       |                               |                                     |                                                |                                                                      |                       |
| • Proteome (LC-MS/MS)                   |                               |                                     |                                                |                                                                      |                       |
| • Genome sequencing                     | • Breast cancer               | • Acquired resistance               | • Biological characterization                  | Navitoclax                                                          | Marczyk et al. (2020) |
| • Methylation (bisulfite sequencing)    |                               |                                     |                                                |                                                                      |                       |
| • ATAC-seq                              |                               |                                     |                                                |                                                                      |                       |
| • Single-cell RNA-seq                   |                               |                                     |                                                |                                                                      |                       |
| • RNA-seq                               |                               |                                     |                                                |                                                                      |                       |
| • ChIP-seq                              | • Breast cancer               | • Acquired resistance               | • Biological characterization                  | Doxorubicin and 5-fluorouracil (5-FU)                                | Mukherjee et al. (2020) |
| • RNA-seq                               |                               |                                     |                                                |                                                                      |                       |
| • Single-cell RNA-seq                   | • Lung cancer                 | • Acquired resistance               | • Biological characterization                  | Cisplatin and paclitaxel                                            | Poojan et al. (2020)  |
| • ATAC-seq                              | • Blood cancer                | • Acquired resistance               | • Biological characterization                  | Bromodomain and Extra-Terminal motif (BET) inhibitor                  | Bell et al. (2019)    |
| • ChIP-seq                              | • Lymphoma                    | • Not specified                      | • Identification of biomarkers                | Anthracycline-based regimen R-CHOP (i.e., rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone) | Fornecker et al. (2019) |
| • RNA-seq                               |                               |                                     |                                                |                                                                      |                       |
| • Proteome (nanoLC-MS/MS)               |                               |                                     |                                                |                                                                      |                       |

(Continued on following page)
TABLE 1 | Recent studies on the use of omics data and computational models to better understand and treat drug-resistant cancer cells.

| Approaches                  | Cancer types          | Resistance type               | Objectives                                    | Drugs                                      | References                     |
|-----------------------------|-----------------------|-------------------------------|-----------------------------------------------|--------------------------------------------|----------------------------------|
| Proteome, phosphoproteome, kinome (LC-MS/MS) | Breast cancer         | Acquired resistance           | Biological characterization                    | 2,5-diaziridinyl-3-hydroxyl-6-methyl-1,4-benzoquinone (RH1) | Kuciauskas et al. (2019) |
| ChiP-seq                    | Breast cancer         | Intrinsic resistance          | Biological characterization                    | Trastuzumab                                | Nasa et al. (2019)              |
| RNA-seq                     | Breast cancer         | Both acquired and intrinsic resistance | Identification of biomarkers                  | Epirubicin, docetaxel, and bevacizumab    | Kim et al. (2018)               |
| Exome sequencing            | Breast cancer         |                               | Biological characterization                    |                                            |                                  |
| Single-cell DNA-seq         | Breast cancer         |                               | Identification of therapeutic targets         |                                            |                                  |
| Single-cell RNA-seq         | Head and neck squamous carcinoma |                               | Biological characterization                    |                                            |                                  |
| Genome sequencing           | Head and neck squamous carcinoma |                               | Biological characterization                    |                                            |                                  |
| mRNA microarray             | Head and neck squamous carcinoma |                               | Biological characterization                    |                                            |                                  |
| Phosphoproteome (LC-MS/MS)  | Head and neck squamous carcinoma |                               | Biological characterization                    |                                            |                                  |
| Network-based modeling      |                       |                               |                                               |                                            |                                  |
| GCNA using mRNA microarray data | Breast cancer         | Not specified                | Identification of biomarkers                  | Trastuzumab and docetaxel                 | Li et al. (2021a)               |
| Cox regression model        | Breast cancer         | Not specified                | Identification of biomarkers                  | Doxorubicin                                | Li et al. (2021b)               |
| Weighted GCNA using RNA-seq data | Breast cancer         | Not specified                | Identification of biomarkers                  | Doxorubicin, cytotoxan, and tamoxifen     | Cui et al. (2020)               |
| Gene co-expression network analysis (GCNA) using RNA-seq data | Breast cancer | Not specified | Identification of biomarkers | Biological characterization |                                            |                                  |
| Methylome (BeadChip array)  | Gastric cancer        | Acquired resistance          | Identification of biomarkers                  | 5-FU and cisplatin                        | Qi and Zhang, (2020)            |
| Genome sequencing           | 19 Types (e.g., Lung cancer, breast cancer, and melanoma) | Not specified | Identification of biomarkers | Biological characterization | 138 Drugs (e.g., vorinostat and bosutinib) | Liu et al. (2019a)             |
| Weighted GCNA using mRNA microarray data | Glioma                | Acquired resistance          | Identification of biomarkers                  | Dibutyryl cyclic adenosine monophosphate  | Zhang et al. (2019)             |
| ceRNA network for correlation between lncRNA and mRNA levels using RNA-seq data | Breast cancer | Not specified | Identification of biomarkers | Biological characterization |                                            |                                  |
| GCNA using RNA-seq data     | Breast cancer         | Acquired resistance          | Identification of biomarkers                  | Docetaxel                                  | Huang et al. (2018)             |
| Cox regression model        | Breast cancer         | Acquired resistance          | Identification of biomarkers                  |                                            |                                  |
| Machine learning            |                       |                               |                                               |                                            |                                  |
| Deep neural network (DNN) with neighborhood component analysis using CNV, somatic mutation, methylome, mRNA microarray, RNA-seq, and proteome data | Breast cancer | Not specified | Prediction of a drug response | 100 Drugs (e.g., tamoxifen) | Malik et al. (2021) |
| Logistic regression using CNV, somatic mutation, mRNA microarray, drug targets, and drug descriptor data | 965 Cell lines from GDSC (lung cancer, urogenital, and leukemia) | Not specified | Prediction of a drug response | 219 Drugs (e.g., AT-7519) for GDSC cell lines | Yu et al. (2021) |
|                            | 491 Cell lines from CCLE | Not specified | Prediction of a drug response | 24 Drugs (e.g., AZD6244) for CCLE cell lines | (Continued on following page) |

(Continued on following page)
TABLE 1 | (Continued) Recent studies on the use of omics data and computational models to better understand and treat drug-resistant cancer cells.

| Approaches                                                                 | Cancer types                  | Resistance type     | Objectives                                      | Drugs                          | References                        |
|----------------------------------------------------------------------------|-------------------------------|--------------------|-------------------------------------------------|-------------------------------|-----------------------------------|
| DNN with multiple elastic nets using mRNA microarray and drug descriptor data | 983 Cell lines from GDSC      | Not specified      | Identification of biomarkers                     | 222 Drugs (e.g., 5-FU) for GDSC cell lines | Choi et al. (2020)                |
|                                                                            | 491 Cell lines from CCLE       |                    | Prediction of a drug response                    | 12 Drugs for CCLE cell lines   |                                   |
| Weighted GCNA, elastic net, and random forest using proteome and phosphoproteome data | NC860 cell line panel         | Not specified      | Identification of biomarkers                     | Various drugs (e.g., cytarabine, 5-FU) | Frejno et al. (2020)              |
| Cox regression model                                                       |                               |                    | Prediction of a drug response                    |                               |                                   |
| Ridge regression and support vector regression using mRNA microarray and RNA-seq data | Colorectal cancer             | Not specified      | Identification of biomarkers                     | 5-FU for colorectal cancer     | Kong et al. (2020)                |
|                                                                            | Bladder cancer                |                    | Prediction of a drug response                    | Cisplatin for bladder cancer   |                                   |
| Ensemble transfer learning (LightGBM, or DNNs with two different architectures) using RNA-seq data and drug descriptor data | Hundreds of cancer cell lines from CCLE, CTRP, gCSI and GDSC | Not specified | Prediction of a drug response                    | Hundreds of drugs from CCLE, CTRP, gCSI and GDSC | Zhu et al. (2020)                |
| Artificial neural network using single-cell metabolome data                 | Leukemia                      | Intrinsic resistance | Prediction of a drug response                    | Cell adhesion as a indication of drug resistance without addition of a drug | Liu et al. (2019b)               |
| DNN using mRNA microarray and RNA-seq data                                | 1,001 Cell lines from 55 tissues (e.g., leukemia) in GDSC | Not specified | Prediction of a drug response                    | Bortezomib, PARP inhibitor, cisplatin, and pacitaxel | Sakellaropoulos et al. (2019) |
| Random forest using RNA-seq, CNV, and methylome data                       | Bladder cancer                | Not specified      | Identification of biomarkers                     | Cisplatin and gemcitabine (for bladder cancer) | Xu et al. (2019)                 |
|                                                                            | Gioma                         |                    | Prediction of a drug response                    | Temozolomide (for glioma)      |                                   |
|                                                                            | Pancreatic cancer             |                    |                                                  | Gemcitabine (for pancreatic cancer) |                                   |
|                                                                            | Gastric cancer                |                    |                                                  | 5-FU (for gastric cancer)      |                                   |
| Elastic net using proteome and kinome data                                 | Colorectal cancer             | Not specified      | Identification of biomarkers                     | 577 Drugs (e.g., cetuximab and afatinib) | Frejno et al. (2017)              |
| Cox regression model                                                       |                               |                    | Prediction of a drug response                    |                               |                                   |

proteins) (Nava et al., 2019; Kagohara et al., 2020; Marczyk et al., 2020; Poojan et al., 2020; Sinkala et al., 2021); and the presence of heterogeneity of a cell population (Niehr et al., 2018), often increasingly examined at a single-cell resolution (Kagohara et al., 2020; Aissa et al., 2021). In a recent study for cell line heterogeneity, for example, application of single-cell DNA and RNA sequencing (RNA-seq) to 20 triple-negative breast cancer (TNBC) patients revealed that rare pre-existing clones having genotypes associated with chemoresistance were adaptively selected in response to neoadjuvant chemotherapy, which subsequently led to acquired transcriptional reprogramming (Kim et al., 2018). For epigenetic alteration, chromosome conformation capture (Hi-C) along with additional omics analyses were conducted for estrogen receptor positive (ER+) breast cancer, which showed that resistance development to endocrine therapy was accompanied with notable 3-dimensional (3D) epigenome alterations (Achinger-Kawecka et al., 2020). Application of multi-omics analyses has also been extended to examine biological processes in quiescent cancer cells that show drug resistance (Lee et al., 2020; Kumar et al., 2021).

Understanding the biology of drug resistance often helps devise effective treatment strategies for drug-resistant cancer cells. Relevant examples (Table 1) include targeting: cancer stem cell phenotypes, in particular stem cell factor receptor c-KIT, for TNBC cells resistant to an anticancer agent RH1 that is currently under clinical trials (Kuciauskas et al., 2019); a range of biological pathways (e.g., metabolism), microenvironment as well as proliferation, migration and invasion of cells, which are all associated with drug resistance for diffuse large B-cell lymphoma patients (Fornecker et al., 2019); zinc finger MYND domain-containing protein 8 (ZMYND8), a putative chromatin reader that appeared to suppress tumorigenic potential and drug resistance induced by doxorubicin (Mukherjee et al., 2020); and EZH2 responsible for histone methylation in taxane-resistant TNBC (Deblois et al., 2020).

As representative examples of overcoming drug resistance on the basis of omics analyses, recent studies additionally conducted
CRISPR-Cas9-based genetic screens to examine cellular plasticity, which was suggested as a therapeutic target for drug-resistant cancer cells (Bell et al., 2019; Torre et al., 2021). Cellular plasticity describes non-genetic state by reprogramming gene expression profiles. In a study by Torre et al., CRISPR-Cas9 genetic screens were implemented for melanoma cells to identify genes that affect cell fate decisions by altering cellular plasticity (Torre et al., 2021).

In particular, modulating the cellular plasticity was demonstrated for vemurafenib inhibiting B-Raf, encoded by a proto-oncogene, in melanoma. Interestingly, inhibiting DOT1L, associated with the onset of melanoma, before the B-Raf inhibition showed more drug resistance than simultaneous inhibition of DOT1 and B-Raf using pinometostat and vemurafenib, respectively. Subsequent transcriptome analysis of knockout cell lines generated clues for non-genetic mechanisms of drug resistance. Another study by Bell et al. focused on acute myeloid leukemia patients that showed non-genetic drug resistance (Bell et al., 2019). Single-cell RNA-seq, followed by CRISPR-Cas9 screening, led to the identification of genes responsible for transcriptional plasticity that triggered epigenetic resistance. Among the genes identified was Lsd1, the inhibition of which was shown to overcome non-genetic drug resistance. As demonstrated by these two recent studies, implementation of genome engineering in addition to omics analyses provides compelling evidence for targets that can help overcome drug resistance.

### COMPUTATIONAL MODELING APPROACHES

While various bioinformatic analyses are available for analyzing omics data, such as enrichment analyses, gene co-expression networks (GCNs) (Cui et al., 2020; Qi and Zhang, 2020) and their variants (e.g., a network of long non-coding RNAs and mRNAs) (Huang et al., 2018; Liu H. et al., 2019) as well as dimensionality reduction (e.g., t-SNE and UMAP), omics data have also been subjected to computational modeling to make predictions for discovering novel mechanisms and devising treatment strategies for drug-resistant cancers (Figure 1). Use of survival analysis in combination with GCNs, and development of a gene regulatory network (GRN) model using a set of ordinary differential equations (ODEs), machine learning models, and genome-scale metabolic models (GEMs) are representative computational modeling approaches that have recently been considered for studying drug-resistant cancer cells (Table 1).

### Network-Based Modeling

GCN has been a popular analysis for understanding gene expression patterns from transcriptome data. GCN is an undirected graph that can be constructed from transcriptome data (e.g., RNA-seq), and connects pairs of genes (nodes in a GCN) with an edge if each pair of genes shows significant co-expression patterns across the transcriptome data. GCN analysis, such as identifying hub genes and/or modules, allows prioritizing candidate genes that may be highly associated with drug resistance of cancer cells. Weighted GCN additionally considers the level of significance in the co-expression relationship between genes in a pair. Often, outcomes from (weighted) GCN analysis are further subjected to other computational analyses, for example survival analysis, to validate the biological and/or clinical significance of the candidate genes. As a recent example, Li et al. focused on **PPP2R2B**, encoding serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B beta isoform, as a potential prognostic biomarker for TNBC on the basis of a series of bioinformatic analyses involving a GCN (Li Z. et al., 2021). Kaplan-Meier survival analysis for this gene revealed that patients with a low expression level of **PPP2R2B** showed shorter survival time than those with a high expression level of **PPP2R2B**. Interestingly, **PPP2R2B** upregulation could attenuate the resistance of TNBC cells to doxorubicin. Likewise, Cox proportional hazards regression model (Cox regression model) was used for genes selected from GCNs to predict prognostic biomarkers for breast cancer, and to suggest genes (e.g., **CCNE2** and **KIF14**) that may help overcome drug resistance (Li Y.-K. et al., 2021).

While GCNs can provide clinically important information when combined with additional predictive models, such as survival analysis above, they have limitations in generating clues on a molecular mechanism associated with development of drug resistance, in particular dynamic interactions between genes. To address this problem, Zhang et al. developed a time-course RNA-seq data-driven computational framework (DryNetMC) to construct GRNs that help elucidate dynamic interactions between genes, and identify key genes associated with mechanisms of drug resistance (Zhang et al., 2019). DryNetMC involves a set of ODEs, a regularized regression method as well as a series of network analyses. Using DryNetMC, GRNs were constructed for dbcAMP-sensitive and dbcAMP-resistant glioma cells based on their time-course RNA-seq data. These differential GRNs were subsequently subjected to a systematic characterization to identify their unique network properties (e.g., node importance) that helped identify key genes (e.g., **KIF2C**, **CCNA2**, **NDC80**, **KIF11**, and **KIF23**) that are predictive of a cancer cell’s drug response. Because network-based models, either by using a GCN or other methods (e.g., ODEs), can visualize a biological context (e.g., association between genes), they will continue to be actively used in the analysis of omics data, and likely along with additional predictive models.

### Machine Learning

Increasing availability of omics data for drug-resistant cancer cells has also provided unprecedented opportunities for building machine learning models. In general, machine learning models perform classification or regression, depending on a given problem. Recently, prediction of anticancer drug response was attempted by using various types of machine learning methods, such as logistic regression (Frejno et al., 2017; Yu et al., 2021), random forest (Xu et al., 2019) and deep neural network (DNN; e.g., multilayer perceptron) (Malik et al., 2021) on the basis of a range of omics and drug response data (Table 1). When developing these machine learning models, transcriptome (RNA-seq or mRNA microarray) was the most frequently adopted dataset, but other types of datasets were also considered, including genome (e.g., gene mutations) (Yu et al., 2021), proteome (Frejno et al., 2020), epigenome (Xu et al.,...
2019), mass spectrometry data (Liu R. et al., 2019) and molecular features of a target drug (Zhu et al., 2020).

In a recent study by Kong et al., a machine learning model was developed that can predict a patient’s drug response on the basis of the analysis of protein-protein interaction (PPI) network and pharmacogenomic data from 3D organoid culture models (Kong et al., 2020). Specifically, potential biomarkers were first inferred from the PPI network analysis, and their corresponding expression profiles along with drug response data (IC50) were used to train a machine learning model (e.g., ridge regression). The resulting drug responses were validated using survival analysis by focusing on colorectal and bladder cancer patients treated with 5-fluorouracil and cisplatin, respectively. The predicted drug responses also appeared to be consistent with transcriptome profiles from drug-sensitive and drug-resistant isogenic cancer cell lines as well as data on somatic mutations associated with already known biomarkers. In this study, consideration of the network analysis not only helped improve the performance of the developed machine learning model, but also facilitated the interpretation of model prediction outcomes. Likewise, in another study, elastic net and random forest regression were used to predict drug responses from abundance data of proteins and their phosphorylation sites in cancer cell lines (Frejno et al., 2020).

Among machine learning methods, DNNs are increasingly used for various predictions, and they have also been used to predict drug responses. Sakellaropoulos et al. developed a DNN model by using GDSC datasets (i.e., transcriptomic data for 1,001 cancer cell lines and IC50 values of 251 drugs) to predict drug responses (Sakellaropoulos et al., 2019). Across several datasets tested, the DNN model showed consistently better performance than elastic net and random forest models. The DNN model was validated by conducting survival analyses for the model-predicted IC50 values, which split patients based on their drug responsiveness. Importantly, pathway enrichment analysis using information from the DNN model (i.e., weights that connect the input layer and the first hidden layer) appeared to associate specific biological pathways with mechanisms of action for drugs. In a more recent study, predicting drug response was also attempted by using a DNN model combined with multiple elastic nets (Choi et al., 2020), referred to as Reference Drug-based Neural Network (RefDNN). RefDNN was developed more in the context of drug resistance, which predicts whether a given cell line is resistant to a target drug by processing gene expression profiles and molecular structure of a drug. RefDNN was also shown to help identify biomarker genes associated with drug resistance, and explore a novel anticancer drug via drug repositioning.

Despite its demonstrated performance, machine learning is often challenged with the limited availability of training datasets for many technical fields. This challenge can be addressed to a certain extent by employing transfer learning as recently demonstrated (Zhu et al., 2020). Zhu et al. demonstrated that ensemble transfer learning can improve the prediction of drug responses in the context of drug repositioning (i.e., use of a drug for another cancer that is already known), precision oncology (i.e., use of a drug for a new cancer that has never been treated before) and new drug development (i.e., use of a new drug for already known cancer). In this particular study, LightGBM (Light Gradient Boosting Machine) and two different DNN models were considered for ensemble transfer learning; larger datasets from the CTRP and GDSC were used as source data for initial training of models, and smaller datasets from CCLE and the Genentech Cell Line Screening Initiative (gCSI) served as target data for further refinement and testing of the models. It was shown that ensemble transfer learning-based models almost always outperformed models that were not developed using transfer learning. This study suggests the use of transfer learning for other drug-resistant cancer cells where a training dataset is sufficiently not available.

**Genome-Scale Metabolic Modeling**

GEM is a computational model that describes gene-protein-reaction (GPR) associations, and can be simulated to predict genome-scale metabolic flux distributions (Gu et al., 2019). GEMs are now available for an increasing number of organisms that are important in biotechnology and biomedicine. Several versions of human GEMs (Ryu et al., 2017; Brunk et al., 2018; Robinson et al., 2020) are currently available, which have been used to examine a target cell’s metabolism, and to predict biomarkers and drug targets for various diseases (Cook and Nielsen, 2017; Gu et al., 2019). For a medical application, a generic human GEM, covering all the known GPR associations in human metabolism, is initially integrated with omics data, often transcriptome data (e.g., RNA-seq), to build a context-specific GEM, a GEM that is specific to a target cell or tissue (Ryu et al., 2015; Opdam et al., 2017). The resulting context-specific GEM is then simulated for various metabolic studies.

Human GEMs have recently been used to study radiation-resistant tumors (Lewis et al., 2021; Lewis and Kemp, 2021), but not drug-resistant cancer cells, to the best of our knowledge. Lewis et al. newly constructed GEMs for radiation-sensitive and radiation-resistant tumors through multi-omics integration (i.e., transcriptome data, mutational data, kinetic data and thermodynamic data) (Lewis et al., 2021). These context-specific GEMs were used to identify changes in redox cofactor production that give resistance to radiation therapy. In the other study, ensemble machine learning classifiers were developed to predict whether an individual is responsive or resistant to a radiation therapy by considering data of metabolite production rates predicted from context-specific GEMs as well as mutation data, transcriptome data and clinical data from TCGA (Lewis and Kemp, 2021). These two studies obviously suggest that GEM-based approaches can also be considered to identify metabolic signatures of drug-resistant cancer cells, and to predict effective drug targets for these cancer cells.

**OUTLOOK**

Understanding genotype-phenotype associations in drug-resistant cancer cells is a highly complex problem, and therefore use of multi-omics data has been considered to capture various aspects of these troubling cancer cells. In particular, multi-omics analyses along with additional tools, such as genome engineering (e.g., CRISPR-Cas9), will continue to play an important role in thorough characterization of drug-resistant cancer cells. Also, an increasing volume of omics data will facilitate development of various types of
computational models. As a consequence, prediction outcomes from computational models will allow more systematically designing experiments for drug-resistant cancer cells.

Despite the promises of omics data and computational models, technical challenges exist. First, current coverage of multi-omics data is not sufficient for thoroughly studying a range of drug-resistant cancer cells. In particular, generation of a consistent set of multi-omics data from each single cell is necessary for in-depth study of a target cancer cell and comparison of different types of cancer cells. Also, it will be interesting to examine the effects of using datasets obtained from patients having a specific disease instead of publicly available datasets (e.g., GDSC and CTRP). While currently available machine learning models have been rigorously validated by using public datasets, they might reveal previously unnoticed limitations in a clinical setting because the public datasets are often generated in a highly controlled condition. In particular, additional consideration of non-genetic factors (e.g., age, gender, and lifestyle) may help reveal new insights on drug-resistant cancer cells. Use of patient-specific datasets will allow more widespread use of the state-of-the-art computational models in a clinical setting.

For network-based modeling, including both GCN and GRN, a breakthrough is needed that allows efficiently developing a cell-specific large-scale GRN that can be simulated under various conditions (e.g., gene perturbation). For machine learning, despite its high predictive performance, there is always a challenge of avoiding overfitting and achieving explainability. Explainability in terms of biological processes is particularly important in the field of biomedicine in order to explain prediction outcomes and make medical decisions. In case of human GEMs, because patient-specific omics data (e.g., RNA-seq) are available to a certain extent, human GEMs should be more actively considered to systematically examine metabolism of drug-resistant cancer cells. Availability of multi-omics data will be particularly useful for interpreting human GEMs and their prediction outcomes; because human GEMs only cover a metabolic network, use of multi-omics data can help explain a complex interplay between metabolic and regulatory networks. Prediction outcomes from the simulation of human GEMs will in turn help explain the insights reaped from omics analyses.

Taken together, advances in omics technologies and computational modeling will bring about positive impacts in understanding and treating drug-resistant cancer cells. Feedback from clinicians and biomedical researchers will be additionally useful for the successful development and clinical application of computational models.

**AUTHOR CONTRIBUTIONS**

HUK, HDJ and YJS wrote the manuscript. HDJ and YJS prepared the figure and table, and critically reviewed the manuscript. All authors contributed to the article and approved the submitted version.

**FUNDING**

This work was supported by the Bio-Synergy Research Project (NRF-2018M3A9C4076475) and the KAIST Cross-Generation Collaborative Lab project of the Ministry of Science and ICT through the National Research Foundation of Korea. This work was also supported by Kwon Oh-Hyun Assistant Professor fund of the KAIST Development Foundation.

**REFERENCES**

Achinger-Kawecka, J., Valdes-Mora, F., Luu, P.-L., Giles, K. A., Caldon, C. E., Qu, W., et al. (2020). Epigenetic Reprogramming at Estrogen-Receptor Binding Sites Alters 3D Chromatin Landscape in Endocrine-Resistant Breast Cancer. *Nat. Commun.* 11 (1), 220. doi:10.1038/s41467-019-14098-x

Aissa, A. F., Islam, A. B. M. K., Ariss, M. M., Go, C. C., Rader, A. E., Conrardy, R. D., et al. (2021). Single-cell Transcriptional Changes Associated with Drug Tolerance and Response to Combination Therapies in Cancer. *Nat. Commun.* 12 (1), 1628. doi:10.1038/s41467-021-21884-z

Bell, C. C., Fennell, K. A., Chan, Y.-C., Rambow, F., Yeung, M. M., Vassiladios, D., et al. (2019). Targeting Enhancer Switching Overcomes Non-genetic Drug Resistance in Acute Myeloid Leukaemia. *Nat. Commun.* 10 (1), 2723. doi:10.1038/s41467-019-10652-9

Brunk, E., Sahoo, S., Zielinski, D. C., Altunkaya, A., Draeger, A., Mih, N., et al. (2018). Recon3D Enables a Three-Dimensional View of Gene Variation in Human Metabolism. *Nat. Biotechnol.* 36 (3), 272–281. doi:10.1038/nbt.4072

Bukowski, K., Kciuk, M., and Kontek, R. (2020). Mechanisms of Multidrug Resistance in Cancer Chemotherapy. *Ijms* 21 (9), 3233. doi:10.3390/ijms21093233

Cai, Q., Wang, S., Jin, L., Weng, M., Zhou, D., Wang, J., et al. (2019). Long Non-coding RNA GBCD-Rlnc1 Induces Chemoresistance of Gallbladder Cancer Cells by Activating Autophagy. *Mol. Cancer* 18 (1), 82. doi:10.1186/s12943-019-1016-0

Choi, J., Park, S., and Ahn, J. (2020). ReDNN: A Reference Drug Based Neural Network for More Accurate Prediction of Anticancer Drug Resistance. *Sci. Rep.* 10 (1), 1861. doi:10.1038/s41598-020-58821-x

Collisson, E. A., Weinstein, J. N., Weinstein, J. N., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., et al. (2013). The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nat. Genet.* 45 (10), 1113–1120. doi:10.1038/ng.2764

Consortium, I. T. P.-C. A. o. W. G. (2020). Pan-cancer Analysis of Whole Genomes. *Nature* 578 (7793), 82–93. doi:10.1038/s41586-020-1969-6

Cook, D. J., and Nielsen, J. (2017). Genome-scale Metabolic Models Applied to Human Health and Disease. *Wires Syst. Biol. Med.* 9 (6), e1393. doi:10.1002/wsbm.1393

Costa, D. B., Nguyen, K.-S. H., Cho, B. C., Sequist, L. V., Jackman, D. M., Riely, G. J., et al. (2008). Effects of Erlotinib in EGFR Mutated Non-small Cell Lung Cancers with Resistance to Gefitinib. *Clin. Cancer Res.* 14 (21), 7060–7067. doi:10.1158/1078-0432.CCR-08-1455

Cui, H., Kong, H., Peng, F., Wang, C., Zhang, D., Tian, J., et al. (2020). Inferences of Individual Drug Response-Related Long Non-coding RNAs Based on Integrating Multi-Omics Data in Breast Cancer. *Mol. Ther. - Nucleic Acids* 20, 128–139. doi:10.1016/j.omtn.2020.01.038

De Luca, A., Parker, L. J., Ang, W. H., Rodolfo, C., Gabbarini, V., Hancock, N. C., et al. (2019). A Structure-Based Mechanism of Cisplatin Resistance Mediated by Glutathione Transferase P1-1. *Proc. Natl. Acad. Sci. USA* 116 (28), 13943–13951. doi:10.1073/pnas.1903297116

Debois, G., Tonekaboni, S. A. M., Grillo, G., Martinez, C., Kao, Y. I., Tai, F., et al. (2020). Epigenetic Switch-Induced Viral Mimicry Evasion in Chemotherapy-Resistant Breast Cancer. *Cancer Discov.* 10 (9), 1312–1329. doi:10.1158/2159-8290.CD-19-1493

Fischer, K. R., Durrans, A., Lee, S., Sheng, J., Li, F., Wong, S. T. C., et al. (2015). Recon3D Enables a Three-Dimensional View of Gene Variation in Human Metabolism. *Nature* 527 (7579), 472–476. doi:10.1038/nature15748
