**Deep transfer learning of cancer drug responses by integrating bulk and single-cell RNA-seq data**

The article addresses the critical challenge of predicting cancer drug responses at the single-cell level, aiming to overcome the limitations posed by cancer cell heterogeneity and varied drug responses among individual cells. The existing research in the field has explored in vitro drug screening studies and leveraged single-cell RNA-sequencing (scRNA-seq) data to uncover the diverse gene expression profiles of cancer subpopulations in response to drugs. However, conventional drug response prediction methods tailored for bulk data are inadequate for the complex and intricate nature of single-cell data, necessitating the development of new computational approaches to infer cancer drug responses accurately.

Deep learning techniques have shown promise in handling large-scale scRNA-seq data, enabling in-depth analysis using advanced neural network architectures. Despite the successes achieved in gene expression imputation and cell clustering with deep learning models, the scarcity of benchmarked single-cell drug response data remains a significant hindrance. In their work, the authors propose a solution called scDEAL (single-cell Drug rEsponse AnaLysis) using a Domain-adaptive Neural Network (DaNN) to predict drug responses from bulk and single-cell RNA-seq data. They leverage deep transfer learning (DTL) to transfer knowledge and relation patterns from bulk data to single-cell data, thereby overcoming the issue of limited training data. Here are the key aspects of their approach:

1. Data Sources: They utilize bulk-level drug response RNA-seq information from the Genomics of Drug Sensitivity in Cancer (GDSC) database and Cancer Cell Line Encyclopedia (CCLE) to train and optimize the model.

2. Data Harmonization: To address differences between bulk and single-cell data structures, scDEAL harmonizes single-cell and bulk embeddings to ensure transferability of drug response labels.

3. Heterogeneity Preservation: To retain heterogeneity in single-cell RNA-seq data, scDEAL includes cell cluster labels for loss function regularization in each training epoch.

4. Interpretability: By employing integrated gradient interpretation, scDEAL infers the signature genes of drug response predictions, enhancing model interpretability.

5. Gene Signatures: They identify gene signatures that directly contribute to drug sensitivity or resistance in a cell by tracing and accumulating the integrated gradients of each neuron in the DTL model.

The findings demonstrate improved accuracy and reliability in predicting how individual cancer cells will respond to specific drugs, compared to using either bulk or single-cell RNA-seq data alone. The study also addresses the potential limitations, such as variations in prediction results based on the collection of bulk gene expression data and sample swaps in certain datasets. By conducting reliability tests, like re-assembling datasets

and optimizing the analysis with scDEAL, the study reaffirms the method's competitive response prediction abilities. The identification of critical gene signatures related to drug responses in re-assembled data underscores the robustness and reliability of scDEAL across different patient datasets, showcasing its potential for broader applications in combined data analysis.

To wrap up, scDEAL offers significant promise for advancing drug development on a single-cell scale. Initially, it enables the anticipation of drug reactions and the connection of gene signatures with treatment impacts. Secondly, the CGs represent promising target signatures suitable for CRISPR screenings or cell reprogramming. Lastly, it can be utilized with current non-drug-treated scRNA-seq data to forecast potential drug reactions across various cell clusters, which can then be chosen for animal drug trials.