Predicting drug response from single-cell expression profiles of tumours

AUTHORS

Simona Pellecchia[1,4,†], Gaetano Viscido[1,2,†], Melania Franchini[1,3], Gennaro Gambardella[1,#]

AFFILIATIONS

[1]Telethon Institute of Genetics and Medicine, Naples, Italy
[2]University of Naples Federico II, Department of Chemical, Materials and Industrial Engineering, Naples, Italy
[3]University of Naples Federico II, Department of Electrical Engineering and Information Technology, Naples, Italy
[4]Scuola Superiore Meridionale, Genomics and Experimental Medicine Program, Naples, Italy

[†]These authors contributed equally to the work.

[#]CORRESPONDENCE: gambardella@tigem.it

ABSTRACT

Drug response prediction at the single cell level is an emerging field of research that aims to improve the efficacy and precision of cancer treatments. Here, we introduce DREEP (Drug Response Estimation from single-cell Expression Profiles), a computational method that leverages publicly available pharmacogenomic screens and functional enrichment analysis to predict single cell drug sensitivity from transcriptomic data. We extensively tested DREEP on several independent single-cell datasets with over 200 cancer cell lines and showed its accuracy and robustness. Additionally, we also applied DREEP to molecularly barcoded breast cancer cells and identified drugs that can selectively target specific cell populations. DREEP provides an in-silico framework to prioritize drugs from single-cell transcriptional profiles of tumours and thus helps in designing personalized treatment strategies and accelerate drug repurposing studies. DREEP is available at https://github.com/gambalab/DREEP.

Introduction

Intra-tumour heterogeneity (ITH) is a recurrent issue in designing effective treatment strategies in clinical practice. It refers to heterogeneous cancer cell populations present in the same tumour that exhibit high cellular plasticity and phenotypic heterogeneity. Accumulating evidence from independent studies shows that ITH gives rise to rare sub-populations of malignant cells characterized by a different epigenetic and transcriptional state that renders them refractory to a variety of anticancer drugs, protecting the whole population from complete eradication (1–4).

Single-cell RNA sequencing (scRNA-seq) has now become a powerful tool to explore ITH at the transcriptional level (5) and thus offers an unprecedented chance to tackle it therapeutically (4, 6, 7). In addition, large-scale pharmacogenomic screens have been performed across hundreds of molecularly characterized cancer cell lines (8–14). The resulting data have provided a valuable resource to link drug sensitivity with genomic and transcriptomic features. Indeed, in the last few years, several machine-learning models have leveraged these datasets to identify molecular markers linked to *in vitro* drug sensitivity in order to predict drug response from -omics profiles of tumour biopsies (8, 15, 16). However, these approaches do not explicitly account for ITH with the result that

genomic and non-genomic biomarkers of drug response have been found for only a restricted number of small molecules. Nonetheless, expression-based biomarkers measured from bulk cell populations have emerged as the most powerful predictors of drug response in vitro for several cytotoxic and targeted drugs, with far more predictive power than other -omics profiles (8, 10). To overcome these limitations, a few recent studies have attempted to leverage scRNA-seq and publicly available drug sensitivity profiles to predict drug response at the single cell level, but they all lack extensive in vitro validation (17, 18).

Here, we developed DREEP (Drug Response Estimation from single-cell Expression Profiles), a bioinformatics tool that leverages results from large-scale cancer cell-line viability screenings with enrichment analysis to predict drug vulnerability from the transcriptional state of a cell. We validate DREEP performance extensively using over 200 cancer cell lines covering 22 distinct cancer types from publicly available single-cell datasets. We also show that DREEP can detect changes in drug sensitivity over time in the breast cancer cell line MCF7 and predict the emergence of drug resistance. Finally, using molecular barcoded MDA-MB-468 cells (4), we show that DREEP can recognize innate drug-tolerant cell subpopulations and predict drug co-treatments that will inhibit all of them. In summary, our results show that it is possible to detect differences in drug response among individual cells within the same tumour. DREEP is implemented in R and is publicly available at https://github.com/gambalab/DREEP.

RESULTS

1. Drug Response Estimation from single-cell Expression Profiles.

DREEP (Drug Response Estimation from single-cell Expression Profiles) is a bioinformatics tool that leverages results from large-scale cell-line viability screens and enrichment analysis to predict drug vulnerability from the transcriptional state of a cell. It requires a *pre-defined* collection of Genomic Profiles of Drug Sensitivity (GPDS); these are lists of genes ranked according to their contribution in correctly predicting the effect of a small molecule. To build GPDS, we integrated several publicly available RNA-seq datasets and cell-line viability screens, including the following datasets: GDSC (8), CTRP2 (13, 19) and PRISM (20) (Figure 1A).

GPDS were computed by correlating the in vitro response to a small molecule from the above studies with the expression of all messenger RNA (mRNA) across all the cancer cell lines for which the drug potency was measured (Figure 1A and methods). In these studies, the potency of a small molecule in a cell line is evaluated as the Area Under the Curve (AUC) of the corresponding dose-response curve. The AUC value reflects cell growth inhibition after 72 hours of treatment, where lower values indicate sensitivity while higher values imply resistance to the tested drug. Thus, a gene whose expression positively correlates with the AUC of a drug across many cell lines can be considered predictive of drug resistance (i.e., the more expressed the gene, the higher the concentration needed to inhibit cell growth) (19). Vice versa, a negatively correlated gene can be considered predictive of drug sensitivity (*i.e.* the more expressed the gene, the lower the concentration needed to inhibit cell growth) (19).

Using this approach, we generated a collection of 2,434 GPDS signatures (i.e. 478 from GDSC, 511 from CTRP, and 1,445 from PRISM) and used them to predict the single-cell-level drug response with the strategy depicted in Figure 1B. Briefly, single cell RNA sequencing data from a sample was first normalized with the gf-icf (Gene Frequency – Inverse Cell Frequency) pipeline to rank genes in each cell according to their relevance (21–23). Then, for each cell we used the most relevant genes as input for Gene Set Enrichment Analysis (GSEA) (24) against each GPDS ranked-list. Since resistance biomarkers are at the top and sensitivity biomarkers are at the bottom of a GPDS ranked lists, a positive value of the enrichment score (ES) returned by DREEP implies that the cell expresses genes associated with drug resistance. Vice versa, a negative ES indicates the cell expresses genes correlated with drug sensitivity. At the end of this process, it is thus possible to estimate the overall effect of each drug on a cell population by estimating the median DREEP enrichment score on that population and compute the percentage of cells sensitive to the drug by simply counting the number of cells in the sample that are associated with a significant ES less than zero.

2. Estimation of the predictive performance of DREEP.

To demonstrate the reliability of the DREEP method and measure its predictive performance, we applied it to the Kinker et al. dataset (25) comprising 53,513 single cell transcriptional profiles from 198 tumorigenic cell lines and spanning 22 distinct cancer types. To evaluate DREEP performance, we first converted predictions from the single-cell level to the cell-line level by computing the median enrichment score for each drug across the cells of each individual cell line. Then, to minimize potential confounders, we evaluated the performance of the DREEP method for each sensitivity dataset independently (Methods). We also searched for the optimal number of relevant genes to use to predict the sensitivity of a cell (*i.e.*, 50, 100, 250, 500, and 1,000 genes). Figure 1C shows the method's overall performance in terms of Precision-Recall and ROC (Receiver Operating Characteristic) curves across 198 cancer cell lines for each drug sensitivity dataset independently (i.e., CTRP2, GDSC and PRISM). In this analysis, predicted drug/cell-line pairs are ranked according to the median enrichment score estimated by DREEP and the performance is evaluated against the corresponding gold standard (Methods). Figure 1D reports the ROC curve's AUC (Area Under the Curve) for all drugs and all cell lines individually. To further test DREEP's performance, we repeated the same analysis, this time applied to a single cell breast cancer dataset (5). This dataset comprises 35,276 individual cells from 32 breast cancer cell lines covering all the main breast cancer tumour subtypes (Luminal/Her2-positive/Basal Like). Figure 1E,F show DREEP's performance in this additional dataset.

These results show that DREEP's performance is significantly better than random, independently of the number of relevant genes used. Thus, as a good compromise between the required computational time and prediction agreement, we decided to use N = 500 genes as a default value for DREEP sensitivity predictions.

Finally, we tested the prediction performance of the DREEP method on both pan-cancer and breast cancer datasets using absolute gene expression levels instead of gf-icf normalization. We observed a drastic reduction of the predictive performance of the method (Supplementary Figure 01) using the absolute gene expression level to estimate its relevance in a cell, showing the importance of data pre-processing prior to the application of a drug prediction algorithm.

3. The DREEP method reconstructs the therapeutic landscape of the pan-cancer dataset.

Next, we used the DREEP method to reconstruct the therapeutic landscape of the 22 cancer types in the cell line pan-cancer dataset of Kinker et al. (25). As shown in Figure 2A (and Supplementary Table 01), we first used DREEP to convert each cell line in a list of small molecules predicted to be effective for at least 90% of profiled cells. We grouped the cells and the drugs into four clusters each, so that cell lines that respond similarly across the tested drugs are grouped together, and analogously drugs that have an effect on the same cell lines are grouped together (i.e. cell line clusters C1, C2 C3, C4 and drug clusters D1, D2, D3 and D4).

These analyses revealed interesting associations between cancer vulnerabilities and drug response. As shown in Figure 2A, cell lines in clusters C1 and C2 comprised 87% (13 out of 15) of head and neck cell lines, 100% of colorectal cell lines (10 out of 10), 83% of bladder cell lines (5 out of 6), 73% of breast cell lines (11 out of 15), 73% of pancreatic cell lines (8 out of 11), and 33% of lung cancer cell lines (13 out of 39). These cancers all have in common Epidermal Grow Factor (EGF) receptor overexpression/amplification, such as EGFR and ERBB2. Indeed, this is one of the major alterations in colorectal (CRC, 60-80%) (26), lung (NSCLC, 40-80%) (27), head and neck (HNSCC, 80-90%) (28, 29), bladder (30, 31), pancreatic (PDAC, 30-89%) (32), and breast cancers (33). For these tumour types, therapies targeting EGFR, ERBB2, or their downstream effectors such as MEK or PI3K, have shown some degree of success (29, 34–36). As depicted in Figure 2A,B, cell-line clusters C1 and C2 exhibit sensitivity to drugs in the D1 cluster that is enriched for drugs whose mode of action is related to the inhibition of EGF receptors (i.e. EGFR, ERBB2), MEK (i.e. MAP2K1, MAP2K2), and PI3K/MTOR (i.e. AKT1, AKT2, AKT3) signalling pathways.

Cluster C3 includes 69% of ovarian cancer cell lines (9 out of 13) and 70% of endometrial/uterine ones (7 out of 10). Cell lines in this cluster are predicted to be sensitive to drugs in the D2 cluster, which is enriched for pro-apoptotic inhibitors (BCL2, BCL2L1, BCL2L2), AURORA kinase inhibitors (AURKA, AURKB), and HDAC inhibitors (HDAC1, HDAC2, HDAC3). Interestingly, ovarian and uterine cancers are usually due to inefficient DNA repair (37). The standard of care for these cancer types is mainly based on neoadjuvant chemotherapy, but several assays on cell lines identified drugs targeting apoptosis or cell cycle regulators as being effective (38, 39, 48–50, 40–47).

Cluster C4 includes all skin cancer cell lines (16 out of 16), all kidney (6 out of 6), all brain (12 out of 12), and 13 lung cancer cell lines. Cell lines in cluster C4 are predicted to be sensitive to multiple families of drugs such as Receptor Tyrosine Kinases (c-MET, KDR, KIT, FLT3, FTL1, FGFR1), BRAF, MTOR, and HDAC1 inhibitors (D3 and D4 drug clusters). This agrees with current clinical practice which uses mTOR and BRAF inhibitors as first-line therapy for kidney cancer and melanoma, respectively (51, 52), since the c-Met/mTOR axis is frequently altered in kidney cancer (53, 54) while BRAF mutations are the main genetic drivers of melanoma (52). Finally, RTKs are frequently altered in glioblastoma (55, 56).

These results show that DREEP can find recurrent drug sensitivity patterns shared by cell lines from multiple cancer types, as well as the relationship between drug response and functional cell status.

4. The DREEP method recapitulates cell population drug response over time.

Next, to evaluate DREEP's ability to recognize adaptation to drug treatment over time, we applied it to the Ben-David et al. dataset (57). As Figure 2C shows, this dataset comprises 7,440 cells from a single cell-derived clone of the MCF7 cell line exposed for two days to 500 nM of bortezomib, a 26S proteasome inhibitor. Specifically, cells were sequenced before treatment (t0), at 12 hrs and 48 hrs following drug exposure (t12, t48), and 96 hrs following drug washout and recovery (t96).

As shown in Figure 2D-E, DREEP predicted that at time t0, t12 and t48, a mixed cell population exists composed of both bortezomib-tolerant and bortezomib-sensitive cells. By contrast, DREEP predicted that at the last timepoint (t96), the majority of cells are resistant to bortezomib. Finally, we were interested in evaluating DREEP's ability to detect distinct drug-response cell subpopulations before and after bortezomib treatment. To this end, we used drug response profiles estimated by DREEP instead of transcriptional profiles to build a new cell-embedding space (Methods). Figure 2F shows that the cell-embedding space built from DREEP sensitivity predictions is better than the mixed cell groups identified using transcriptional profiles (Figure 2C) at separating MCF7 cells into clusters resembling the drug exposure over time.

In summary, these analyses demonstrate DREEP's ability to recapitulate drug response and variability over time.

4. The DREEP method identifies small molecules that selectively inhibit cell subpopulations.

Finally, we evaluated DREEP's ability to identify innate drug-resistant cell populations and to suggest drug co-treatments. To this end, we used single cell data of the untreated barcoded MDA-MB-468 breast cancer cell line that we recently generated (4). This dataset comprises 1,541 single cells, each associated with a molecular lineage (i.e. barcode) that have been shown to be either tolerant or sensitive to afatinib, a tyrosine kinase inhibitor of EGFR and HER2 receptors (4) (Figure 3A).

We applied DREEP to each of the two MDAMB468 cell subpopulations (i.e. afatinib-tolerant and afatinib-sensitive) and predicted drugs able to selectively inhibit the growth of either. We thus found 156 drugs predicted to preferentially inhibit the afatinib-tolerant subpopulation and 80 drugs (FDR<1% and Supplementary Table 02) for the afatinib-sensitive subpopulation (Figure 3B), which included 11 out of 12 EGFR inhibitors (Supplementary Table 02).

Interestingly, the most overrepresented classes among the drugs predicted to selectively inhibit the afatinib-tolerant subpopulation were IGF1R, HDAC, and AKT/MTOR/PI3K inhibitors (Figure 3C). We have already shown that afatinib and IGF1R inhibitors have a strong synergistic effect on these cells (4), thus validating the DREEP predictions that IGF1R inhibitors should deplete the cell population that do not respond to afatinib in this cell line.

To experimentally validate these predictions further, we selected three HDAC inhibitors (Tenovin-6, bellinostat, and AR-42) and three AKT/MTOR/PI3K inhibitors (Uprosertib, PIK93, and buparlisib) from the list of drugs predicted to selectively inhibit the afatinib-tolerant subpopulation. We then measured experimentally if these drugs had an additive, synergistic, or antagonistic effect on MDA-MB-468 cells in combination with afatinib (Supplementary Table 03). As shown in Figure 3D, the median synergy score of each of the six selected drugs tested in combination with afatinib is compatible with an additive effect (58), suggesting that the two drugs work independently on different subpopulations of cells. Finally, as Figure 3E shows, when cells are embedded into an embedding space reconstructed from DREEP sensitivity profiles (Methods), two main cell clusters appear that are enriched for either afatinib-tolerant or afatinib-sensitive lineages, in contrast to the mixed cell groups identified using transcriptional profiles (Figure 3A). These results show how the DREEP method alone can recognise that there are two different cell populations in this cell line characterized by a distinct drug response profile.

Altogether, these results show that DREEP can predict drug sensitivity at the single cell level and identify drugs able to selectively inhibit a subpopulation of cells that can be then combined to find drug combinations.

Discussion

Intra-tumour heterogeneity (ITH) presents a significant challenge in developing effective cancer treatments and achieving precision oncology. ITH refers to the variation of tumour cells both within and between tumours. As such, it is crucial to have methods to accurately identify the drug sensitivities of different tumour clones and determine the most effective treatment for each cancer type and the individual patient. To address this challenge, we introduced DREEP, a novel computational method that predicts drug sensitivity at the single cell level and can propose cell population-specific treatments. To achieve this, DREEP uses a set of drug-sensitivity signatures obtained by correlating gene expression and drug sensitivity profiles across hundreds of tumorigenic cell lines. It then uses these signatures to compute an enrichment score reflecting the degree of sensitivity of a cell to a specific drug.

To test the predictive performance and reliability of DREEP, we applied it to two independent datasets: (i) the pan-cancer cell line dataset from Kinker et al. comprising 198 tumorigenic cell lines from 22 distinct cancer types (25) and (ii) the single cell breast cancer atlas from Gambardella et al. that includes 32 breast cancer cell lines (5). We observed that DREEP successfully identifies the most effective drugs for each cell line with an overall performance that is several folds better than random on both datasets. We also successfully used DREEP to reconstruct the single-cell variability in drug response across the 198 cell lines of the pan-cancer dataset and to identify the functional cellular activities linked to common drug response patterns among cancer types and cell lines. In addition, when used to analyse single-cell data collected from the breast cancer cell line MCF7 exposed to bortezomib over time (57), DREEP easily reconstructed the interchange between the sensitive and tolerant cell populations. Finally, using single cell data of the molecularly labelled MDA-MB-468 cell line (4), we showed that DREEP can identify innate drug-resistant cell populations and predict drug co-treatments able to deplete these cell populations. Overall, we show that DREEP produces an accurate reconstruction of the variability of drug response in cancer cell lines from single cell

transcriptomics. The DREEP R package is available at https://github.com/gambalab/DREEP.

In the future, it would be crucial to evaluate how the integration of other single-cell -omics measurements, such as genomic, proteomic, and epigenetic cell states, can improve the identification of variations in drug responses among cancer cell populations. Moreover, since tumour heterogeneity is linked to clonal evolution processes, it would be also important to consider the dynamics of clonal populations under therapy pressure to improve the performance of future computational methods that will be developed for personalized diagnosis and treatment of cancer patients.

Figure & Legends

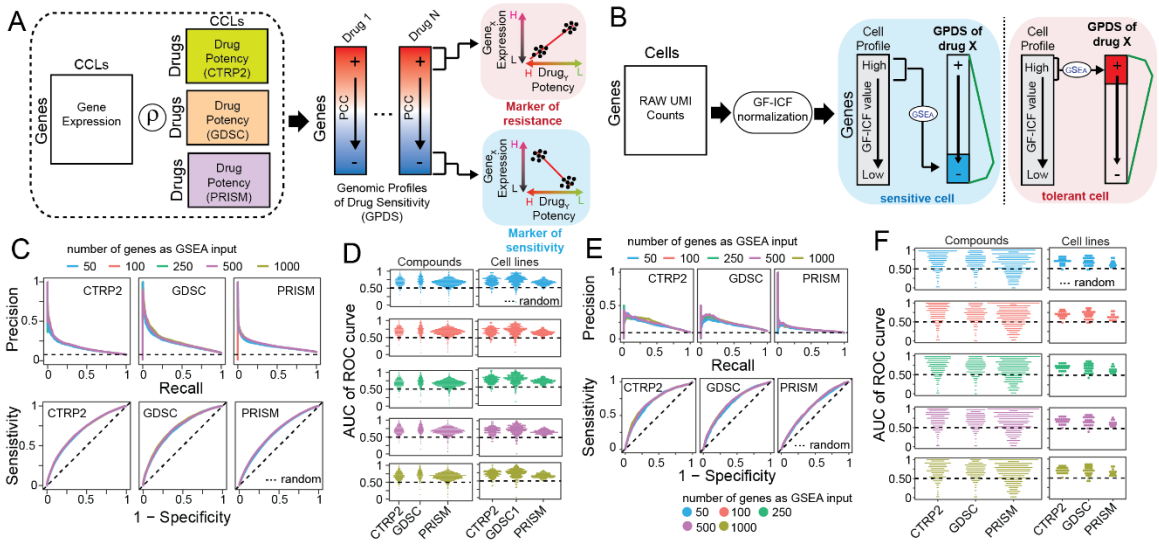


Figure 1 – Drug Response Estimation from single-cell Expression Profiles (DREEP). (**A**) Construction of Genomic Profiles of Drug Sensitivity (GPDS). The potency profile of each drug in terms of AUC (Area Under the Curve) is correlated with the expression of all messenger RNA across the cell lines for which the drug potency was measured. (**B**) Bioinformatics pipeline for the prediction of single-cell drug response. Raw UMI counts are first normalized with the gf-icf pipeline and then the most relevant genes in a single cell are used as input for a GSEA against the GPDS ranked-list to predict its drug sensitivity. (**C**) Performance of DREEP drug sensitivity on 198 cell lines for each set of GPDS signatures used and different numbers of relevant genes to predict the effect of a drug. DREEP's performance is estimated using either Precision-recall curve (upper) or ROC curve (bottom). (**D**) Left column: Each point is a drug whose y-coordinate is AUC of the corresponding ROC curve across the 198 cell lines and computed by sorting cell lines from the most to the least sensitive to the drug. Each panel reports the results using different numbers of relevant genes to predict the effect of the drug. Right column: Each point is a cell line whose y-coordinate is the AUC of the corresponding ROC curve obtained by sorting drug predictions from the most to the least sensitive. Each panel reports the results using different numbers of relevant genes to predict the effect of a drug. (**E**) Same as (C) but for the 32 breast cancer cell lines published in Gambardella et al. (**F**) Same as (D) but for 32 breast cancer cell lines published in Gambardella et al.

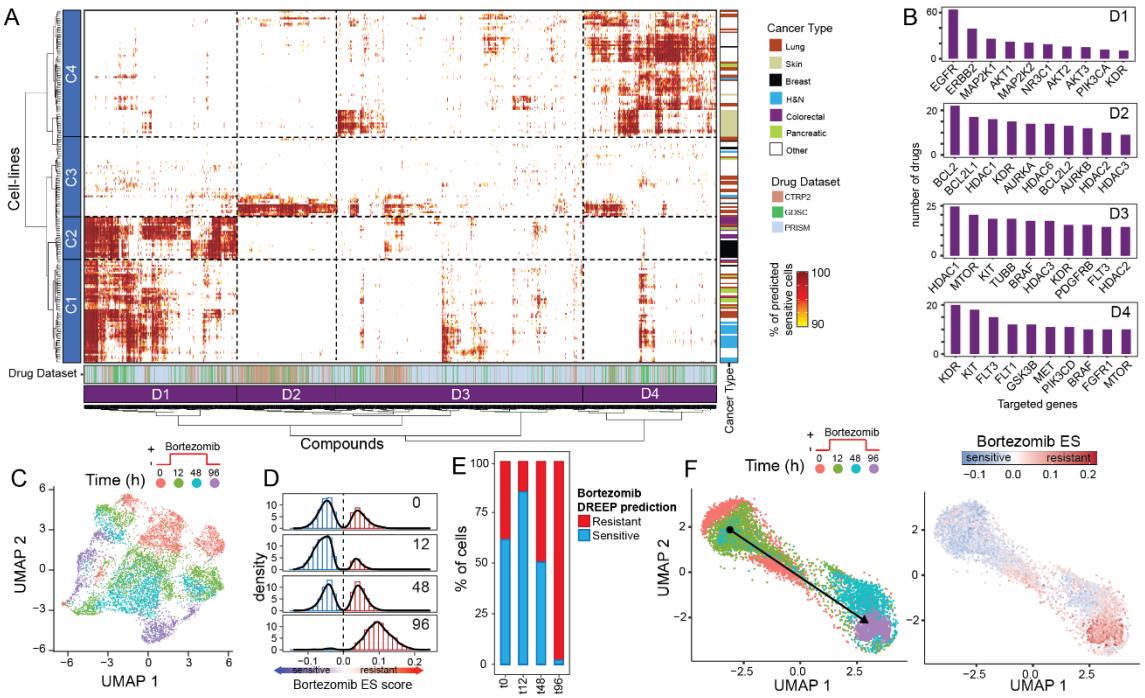


Figure 2 – DREEP captures single cell variability in drug response. (**A**) Heatmap depicting both cell-line clusters (rows) and cancer-specific drug clusters (columns). Cell lines are colour-coded according to the cancer type while drugs are colour-coded according to the dataset to which they belong. Both rows and columns of the heatmap were grouped in four clusters named C1-C4 for cell line (rows) and D1-D4 for drugs (columns). (**B**) Bar-plot depicting the number of drugs sharing the same target in each cancer-specific drug cluster. Only the top ten target genes for each cluster are shown. (**C**) UMAP plot of 7,440 cells of the MCF7 cell line exposed to 500 nM bortezomib. (**D**) Bortezomib DREEP enrichment score (ES) distribution at each timepoint for the 7,440 cells of the MCF7 cell line in (C). A positive or negative score means that the cell is predicted to be resistant or sensitive to bortezomib, respectively. (**E**) Percentage of predicted bortezomib sensitive or tolerant cells for each time point. (**F**) UMAP plot of the same cells in (C) but using the drug profile estimated by DREEP instead of its corresponding transcriptional profile for each cell. In the left panel, cells are colour-coded according to the sequencing timepoint, while in the right panel cells are colour-coded according to the bortezomib score predicted by DREEP. A positive or negative score means that the cell is predicted to be resistant or sensitive to bortezomib, respectively.

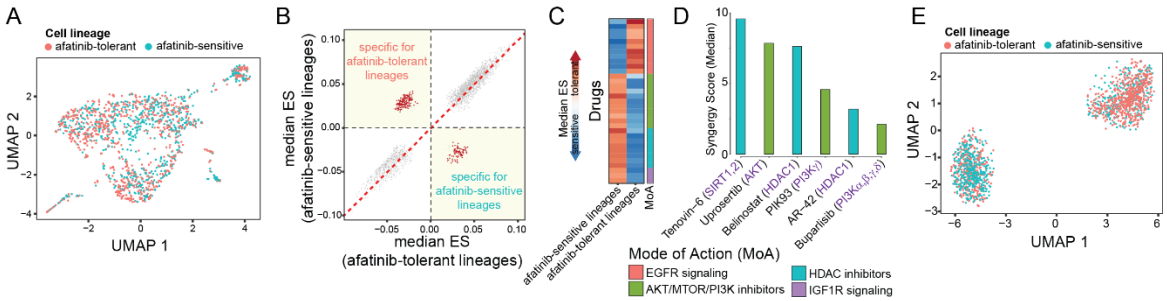


Figure 3 – DREEP identifies drugs that can selectively inhibit a subpopulation of cells. (**A**) UMAP plot of 1,541 cells of MDA-MB-468 cells from Pellecchia et al. (4). Cells are colour-coded according to whether they belong to an afatinib-sensitive or tolerant lineage. (**B**) Each point represents a drug whose coordinates are the median enrichment score predicted by DREEP for the drug in afatinib-tolerant cells [x-axis] and afatinib-sensitive cells [y-axis]. The more negative the value of the enrichment score, the more potent the predicted effect of the drug. The drugs specifically inhibiting at least one of the two subpopulations are indicated in red. (**C**) Heatmap showing enriched drug classes predicted to specifically inhibit afatinib-tolerant or afatinib-sensitive cells. Each row of the heatmap represents a drug coloured according to its median enrichment score predicted by DREEP across cells of the same lineage (i.e. afatinib-tolerant or afatinib-sensitive). (**D**) Median synergy score of drugs predicted by DREEP to selectively inhibit afatinib-tolerant lineages in MDA-MB-468 cells. Each drug was tested in combination with afatinib, and the synergy score computed using the Loewe statistical model. A positive score represents a synergism between the two tested drugs. (**E**) UMAP plot of the cells in (A) but using the drug profile estimated by DREEP instead of the corresponding transcriptional profile for each cell. Cells are colour-coded according to whether they belong to an afatinib-sensitive or tolerant lineage.

Methods

scRNA-seq datasets processing: The scRNA-seq data of the cell line pan-cancer dataset from Kinker et al. (25) was obtained from the Broad Institute's single cell portal (SCP542). Specifically, CPM count matrix was downloaded. Before being processed, gene symbols were converted into ensemble id and only ensemble id associated with no more than one gene were retained. At the end of this pre-processing step, the final cell line pan-cancer matrix contained 16,558 genes and 53,513 cells. Raw scRNA-seq data of the MCF7 cell line exposed to bortezomib (57) was obtained from GEO (GSE114461) and only cells with at least 5,000 UMI were retained. The single cell breast cancer atlas dataset was obtained from figshare at the following address https://doi.org/10.6084/m9.figshare.15022698.

Bulk RNA-seq dataset processing: Raw counts and associated metadata of the bulk RNA-seq data of cancer cell lines used to build GPDS signature were obtained from the depmap portal (https://depmap.org). Cell lines associated with haematopoietic tumours were excluded. Raw counts were normalised with edgeR (59) and transformed into log10 (CPM+1). Poorly expressed genes and genes whose entropy values were in the fifth percentile were discarded. Entropy values were estimated by discretizing the expression of each gene in ten equally spaced bins using the functions included in the R package named *entropy*. After these filtering steps, the final expression matrix comprised 1,117 cell lines and 13,849 genes.

Drug sensitivity dataset processing: The drug sensitivity dataset of CTRP2 (Cancer Therapeutic Response Portal), GDSC (Genomics of Drug Sensitivity in Cancer) and PRISM (Profiling Relative Inhibition in Mixtures) were all obtained from the depmap portal (https://depmap.org). Only drugs for which the in vitro response was available for at least 100 cell lines out of the 1,117 for which gene expression data was available were retained. After this filtering, the final drug sensitivity dataset comprised 478 drugs from the GDSC dataset, 511 drugs from the CTRP2 dataset, and 1,445 drugs from the PRISM dataset. For all three cell viability datasets, drug potency was expressed in terms of AUC (Area Under the Curve) of the corresponding dose response graph.

Construction of GPDS ranked-lists: Genomic Profiles of Drug Sensitivity (GPDS) ranked-lists were built by computing the Pearson correlation coefficient (PCC) between the expression of each gene and the

potency of a drug across the cell lines for which drug potency was measured. At the end of this process, each GPDS is a ranked list of 13,849 expression-based biomarker genes whose degree of PCC reflects their importance in predicting the effect of a small molecule. This is because the AUC value of a drug reflects its cell line growth inhibition, with lower values indicating sensitivity and higher values implying tolerance to the drug. Thus, a gene whose expression positively correlates with the AUC of a drug is predictive of resistance to that drug, conversely a negatively correlated gene is predictive of sensitivity to the drug.

Drug Response Estimation from single-cell Expression Profiles (DREEP): To predict drug response at the single-cell level, we first used gf-icf normalization (https://github.com/gambalab/gficf) (22) to extract the top relevant genes from each cell. We then used these gene sets as input for Gene Set Enrichment Analysis (GSEA) (24) against each GDPS ranked-list to predict the sensitivity of a cell to a specific drug. Since GPDS ranked-lists contain predictive biomarkers of resistance at the top and predictive biomarkers of sensitivity at the bottom for each drug, a positive enrichment score implies that the cell expresses genes that render it tolerant to the drug, while a negative enrichment score indicates the cell expresses genes that render it sensitive to the drug. Finally, by estimating the median enrichment score on a cell population and counting the number of cells with a significant enrichment score less than zero, it is possible to estimate the overall effect of a drug on that population and the percentage of cells sensitive to it. The DREEP package is implemented in R and it is available on github at the following address https://github.com/gambalab/DREEP. Enrichment scores and associated statistics are computed using the *fgsea* package (60).

Estimation of DREEP performances: To estimate DREEP's precision and sensitivity in predicting drug response from scRNA-seq data of cancer cell lines, we used the three cell-viability screening datasets from CTRP2, GDSC and PRISM described above. For each cell-viability screening dataset, the corresponding "gold standard" is composed of pairs of cell lines and drugs for which the corresponding cell line is sensitive. To build the "gold standard" of each cell-viability screening dataset, we first computed the z-score percentiles from the AUC of each drug across the cell lines for which AUC was available and then defined a cell line sensitive to a drug only if its z-score was in the 5% percentile. Next, we applied DREEP to scRNA-seq data of each cell line and computed its median enrichment score for each drug. Finally, Precision-Recall and ROC curves on each cell-viability screening dataset depicted in Figure 1C-F were computed with the *precrec* R package (https://github.com/evalclass/precrec) using the corresponding "gold standard" and sorting DREEP drug/cell lines pairs according to the estimated median cell line enrichment score multiplied by minus one in order to have drugs predicted to be sensitive at the top.

Cell embedding space from DREEP predictions: DREEP output can be arranged in a drug response matrix E of dimensions $n \times m$ where n is equal to the number of cells and m equal to the number of GPDS used. Each element $E_{i,j}$ of matrix E ranges within the $[-1, +1]$ interval and it is different from zero only if the enrichment score between the relevant genes of the cell i and the GPDS ranked-list j is significant (i.e. P<0.05). Specifically, $E_{i,j} > 0$ indicates the cell i is predicted resistant to the drug j, while $E_{i,j} < 0$ indicates the cell i is predicted sensitive to the drug j, and $E_{i,j} = 0$ indicates that there is no supporting evidence to decide if cell i is sensitive or not to the drug j. Next, we computed how much two cells have a similar predicted drug response profile as the distance between two cells x and y with *Fuzzy Jaccard Distance* defined as follow:

$$FJD_{x,y} = 1 - \frac{\sum_{j=1}^{m} \left| E_{x,j} + E_{y,j} \right|}{2 * m}$$

FJD is a non-negative symmetric distance matrix of dimensions $n \times n$ whose elements range within the [0,1] interval. Specifically, each element $FJD_{x,y}$ of FJD quantifies how many of the drugs to which the cell x is sensitive are shared with the cell y and how many of the drugs the cell x is resistant to being shared with the cell y. The more two cells share drugs that are predicted to have the same effect, the closer to 0 is the $FJD_{x,y}$ value, and vice versa for two cells with opposite drug predictions, $FJD_{x,y}$ equals 1. It can be easily shown that the Fuzzy Jaccard Distance we implemented satisfies all the conditions to be defined as a metric, including the triangle inequality constraint. Finally, the cell embedding space was built using the precomputed cell-to-cell distance matrix FJD as input of the *umap* function of the uwot R package

(https://github.com/jlmelville/uwot). Cell embedding from DREEP predictions has been implemented in the function *runDrugReduction* of the DREEP package.

The therapeutic landscape of the cell line pan-cancer dataset (Figure 2A): We first used DREEP to predict the effect of the 2,434 drugs on each of the 53,513 single-cell transcriptional profiles of Kinker et al. dataset (25). Then, we converted predictions from the single-cell level to the cell-line level by counting for each drug the fraction of cells predicted to be sensitive (i.e. DREEP enrichment score < 0 and P<0.05). At the end of this process, we obtained a matrix $A_{198\times2,434}$ where $A_{i,j}$ contains the percentage of cells of the cell line i predicted to be sensitive to the drug j. Next, we set elements of A less than 0.9 as being equal to 0 while all elements of A greater or equal to 0.9 were set as equal to 1. Then, we hierarchically clustered rows (i.e. cell lines) and columns (i.e. drugs) of A using the Jaccard distance. Jaccard distances between row pairs and column pairs of the matrix A were computed using the *dist* function of the R statistical environment. Clustering of both, cell lines and drugs, was performed using the *hclust* function and resulting dendrograms cut with the *cutree* function of R statistical environment.

Differential drug sensitivity analysis: We used DREEP to predict the effects of 2,434 drugs on 1,541 sequenced cells of the barcoded MDA-MB-468 cell line (4). Then, we performed a Mann-Whitney test and corrected p-values using the Benjamini-Hochberg method to identify drugs with differential sensitivity between afatinib-sensitive and afatinib-tolerant subpopulations. Finally, a drug was considered specific for a subpopulation if its FDR was less than 0.1 and the median enrichment score across the other subpopulation was greater than zero. This analysis has been implemented in the *runDiffDrugAnalysis* function of the DREEP package.

Drug combinations assay: 4 x 10^4 MDA-MB-468 cells were seeded in a 96-well plate and incubated overnight at 37 °C. The two drugs to test were prepared in different dilutions and then combined in all possible drug pairs to generate a 5 x 5 combination matrix. Then, cells were exposed to either single agent drug or to the drug pairs of the drug combination matrix, while negative controls were treated with DMSO (each treatment was performed in triplicate). Following 72 hr incubation at 37 °C, cell viability was measured with CellTiter (Promega) and the absorbance was read at 490 nm with the plate reader GloMax® Discover instrument. The drug interactions and expected drug responses were calculated with the Combenefit tool (58) using the Loewe additivity model.

Code availability: The DREEP package is available on github at the following address https://github.com/gambalab/DREEP.

Data availability: All the processed datasets and associated metadata, including scRNA-seq, bulk RNA-seq, and sensitivity profiles used in this work have been deposited on figshare at the following address https://doi.org/10.6084/m9.figshare.23261234.

Funding: This work was supported by the My First AIRC grant 23162.

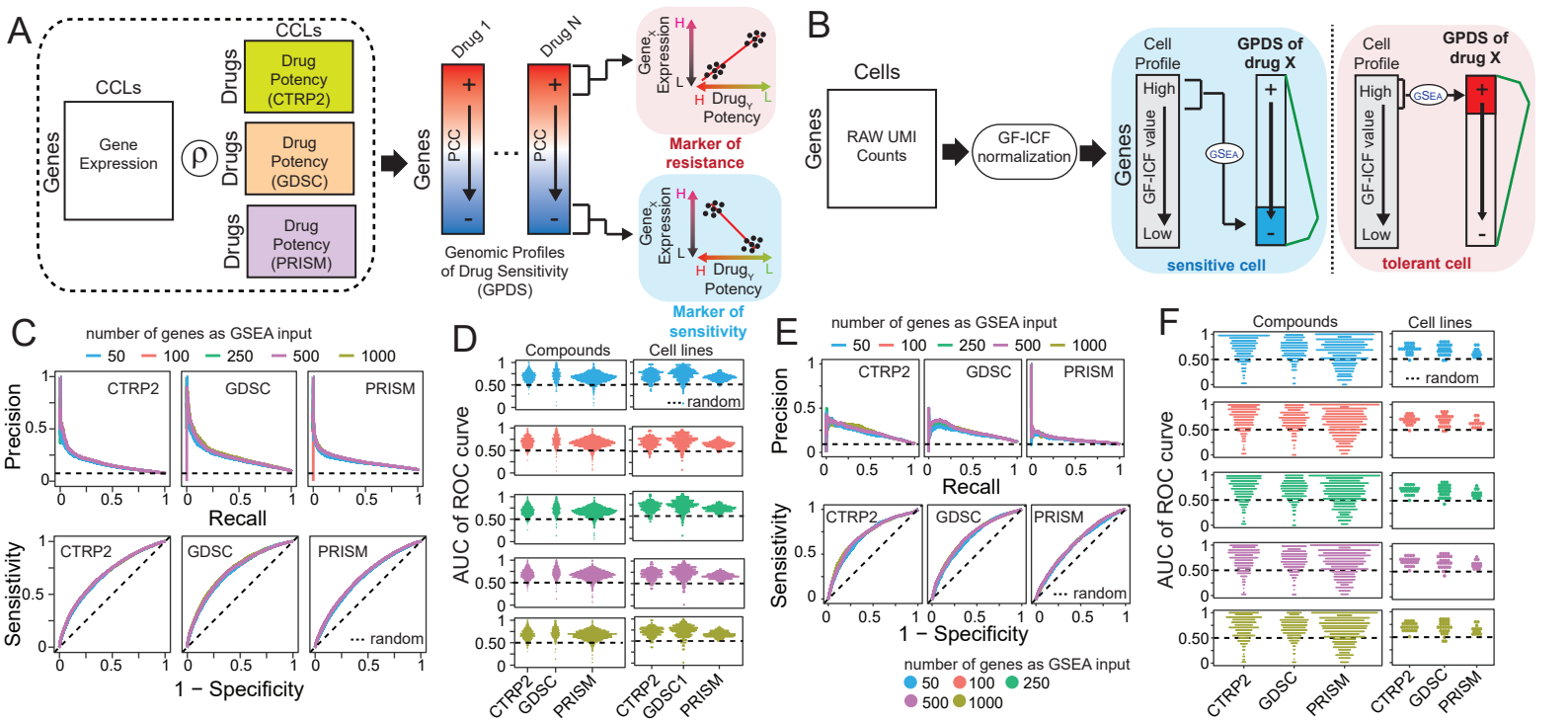
Author Contribution: SP and GV performed drug combination experiments. MF helped with bioinformatics analysis and in writing the manuscript. GG supervised the work, wrote the manuscript, and conceived the original idea.

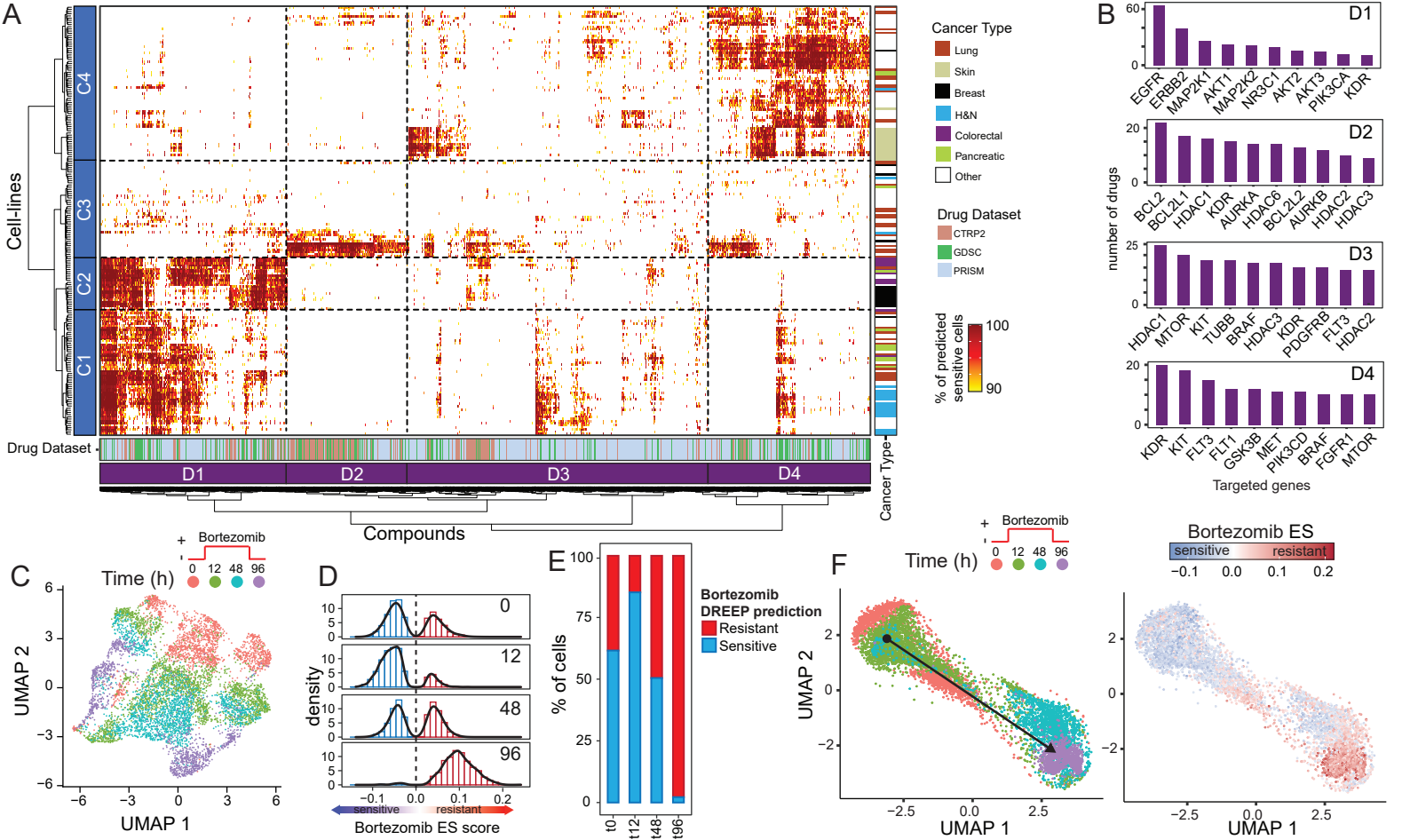
Competing interests: The authors declare no competing interests.

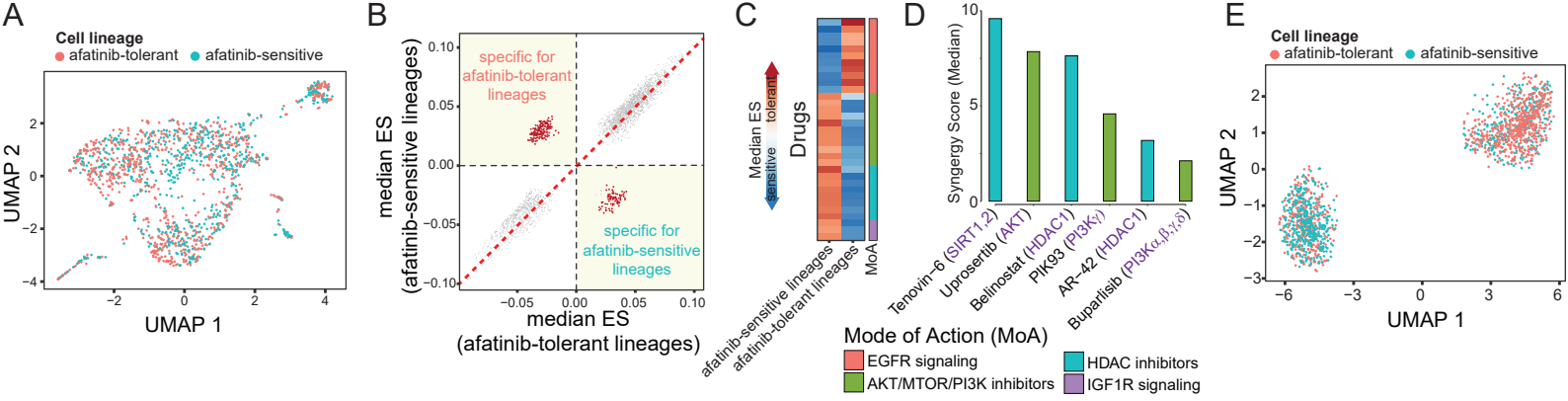
References

1. Shaffer,S.M., Dunagin,M.C., Torborg,S.R., Torre,E.A., Emert,B., Krepler,C., Beqiri,M., Sproesser,K., Brafford,P.A., Xiao,M., *et al.* (2017) Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature*, **546**, 431–435.
2. Sharma,S. V., Lee,D.Y., Li,B., Quinlan,M.P., Takahashi,F., Maheswaran,S., McDermott,U., Azizian,N., Zou,L., Fischbach,M.A., *et al.* (2010) A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. *Cell*, **141**, 69–80.
3. Tirosh,I., Izar,B., Prakadan,S.M., Wadsworth,M.H., Treacy,D., Trombetta,J.J., Rotem,A., Rodman,C., Lian,C., Murphy,G., *et al.* (2016) Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (80-.).*, **352**, 189–196.
4. Pellecchia,S., Franchini,M., Viscido,G., Arnese,R. and Gambardella,G. (2023) Single cell lineage tracing reveals subclonal dynamics of anti-EGFR therapy resistance in triple negative breast cancer. *bioRxiv*, 10.1101/2023.04.04.535588.
5. Gambardella,G., Viscido,G., Tumaini,B., Isacchi,A., Bosotti,R. and di Bernardo,D. (2022) A single-cell analysis of breast cancer cell lines to study tumour heterogeneity and drug response. *Nat. Commun.*, **13**, 1714.
6. Li,X., Francies,H.E., Secrier,M., Perner,J., Miremadi,A., Galeano-Dalmau,N., Barendt,W.J., Letchford,L., Leyden,G.M., Goffin,E.K., *et al.* (2018) Organoid cultures recapitulate esophageal adenocarcinoma heterogeneity providing a model for clonality studies and precision therapeutics. *Nat. Commun.*, **9**, 2983.
7. Saeed,K., Ojamies,P., Pellinen,T., Eldfors,S., Turkki,R., Lundin,J., Järvinen,P., Nisen,H., Taari,K., af Hällström,T.M., *et al.* (2019) Clonal heterogeneity influences drug responsiveness in renal cancer assessed by ex vivo drug testing of multiple patient-derived cancer cells. *Int. J. Cancer*, **144**, 1356–1366.
8. Iorio,F., Knijnenburg,T.A., Vis,D.J., Bignell,G.R., Menden,M.P., Schubert,M., Aben,N., Gon??alves,E., Barthorpe,S., Lightfoot,H., *et al.* (2016) A Landscape of Pharmacogenomic Interactions in Cancer. *Cell*, **166**, 740–754.
9. 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.
10. Costello,J.C., Heiser,L.M., Georgii,E., Gonen,M., Menden,M.P., Wang,N.J., Bansal,M., Ammad-ud-din,M., Hintsanen,P., Khan,S.A., *et al.* (2014) A community effort to assess and improve drug sensitivity prediction algorithms. *Nat Biotechnol*, **32**, 1202–1212.
11. Daemen,A., Griffith,O.L., Heiser,L.M., Wang,N.J., Enache,O.M., Sanborn,Z., Pepin,F., Durinck,S., Korkola,J.E., Griffith,M., *et al.* (2013) Modeling precision treatment of breast cancer. *Genome Biol*, **14**, R110.
12. Heiser,L.M., Sadanandam,A., Kuo,W.L., Benz,S.C., Goldstein,T.C., Ng,S., Gibb,W.J., Wang,N.J., Ziyad,S., Tong,F., *et al.* (2012) Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proc Natl Acad Sci U S A*, **109**, 2724–2729.
13. 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.* (2015) Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat. Chem. Biol.*, **12**, 1–10.
14. Trastulla,L., Noorbakhsh,J., Vazquez,F., McFarland,J. and Iorio,F. (2022) Computational estimation of quality and clinical relevance of cancer cell lines. *Mol. Syst. Biol.*, **18**, e11017.
15. Wu,L., Wen,Y., Leng,D., Zhang,Q., Dai,C., Wang,Z., Liu,Z., Yan,B., Zhang,Y., Wang,J., *et al.* (2022) Machine learning methods, databases and tools for drug combination prediction. *Brief. Bioinform.*, **23**, bbab355.
16. Adam,G., Rampášek,L., Safikhani,Z., Smirnov,P., Haibe-Kains,B. and Goldenberg,A. (2020) Machine learning approaches to drug response prediction: challenges and recent progress. *npj Precis. Oncol.*, **4**, 19.
17. Fustero-Torre,C., Jiménez-Santos,M.J., García-Martín,S., Carretero-Puche,C., García-Jimeno,L., Ivanchuk,V., Di Domenico,T., Gómez-López,G. and Al-Shahrour,F. (2021) Beyondcell: targeting cancer therapeutic heterogeneity in single-cell RNA-seq data. *Genome Med.*, **13**, 187.
18. Suphavilai,C., Chia,S., Sharma,A., Tu,L., Da Silva,R.P., Mongia,A., DasGupta,R. and Nagarajan,N. (2021) Predicting heterogeneity in clone-specific therapeutic vulnerabilities using single-cell transcriptomic signatures. *Genome Med.*, **13**, 189.
19. Rees,M.G., Brenan,L., do Carmo,M., Duggan,P., Bajrami,B., Arciprete,M., Boghossian,A., Vaimberg,E., Ferrara,S.J., Lewis,T.A., *et al.* (2022) Systematic identification of biomarker-driven drug combinations to overcome resistance. *Nat. Chem. Biol.*, **18**, 615–624.
20. Corsello,S.M., Nagari,R.T., Spangler,R.D., Rossen,J., Kocak,M., Bryan,J.G., Humeidi,R., Peck,D., Wu,X., Tang,A.A., *et al.* (2020) Discovering the anticancer potential of non-oncology drugs by systematic viability profiling. *Nat. Cancer*, **1**, 235–248.
21. Gambardella,G. and di Bernardo,D. (2019) A Tool for Visualization and Analysis of Single-Cell RNA-Seq Data Based on Text Mining. *Front. Genet.*, **10**.
22. Franchini,M., Pellecchia,S., Viscido,G. and Gambardella,G. (2023) Single-cell gene set enrichment analysis and transfer learning for functional annotation of scRNA-seq data. *NAR Genomics Bioinforma.*, **5**, lqad024.
23. Slovin,S., Carissimo,A., Panariello,F., Grimaldi,A., Bouché,V., Gambardella,G. and Cacchiarelli,D. (2021) Single-Cell RNA Sequencing Analysis: A Step-by-Step Overview BT - RNA Bioinformatics. In Picardi,E. (ed). Springer US, New York, NY, pp. 343–365.
24. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S., *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 15545–50.
25. Kinker,G.S., Greenwald,A.C., Tal,R., Orlova,Z., Cuoco,M.S., McFarland,J.M., Warren,A., Rodman,C., Roth,J.A., Bender,S.A., *et al.* (2020) Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat. Genet.*, **52**, 1208–1218.
26. Pabla,B., Bissonnette,M. and Konda,V.J. (2015) Colon cancer and the epidermal growth factor receptor: Current treatment paradigms, the importance of diet, and the role of chemoprevention. *World J. Clin. Oncol.*, **6**, 133–141.
27. Prabhakar,C.N. (2015) Epidermal growth factor receptor in non-small cell lung cancer. *Transl. Lung Cancer Res. Vol 4, No 2 (April 19, 2015) Transl. Lung Cancer Res. (Molecular Genet. Lung Cancer)¹*
28. Stransky,N., Egloff,A.M., Tward,A.D., Kostic,A.D., Cibulskis,K., Sivachenko,A., Kryukov,G. V, Lawrence,M.S., Sougnez,C., McKenna,A., *et al.* (2011) The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. *Science (80-.).*, **333**, 1157–1160.
29. Lawrence,M.S., Sougnez,C., Lichtenstein,L., Cibulskis,K., Lander,E., Gabriel,S.B., Getz,G., Ally,A., Balasundaram,M., Birol,I., *et al.* (2015) Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*, **517**, 576–582.

30. Chaux,A., Cohen,J.S., Schultz,L., Albadine,R., Jadallah,S., Murphy,K.M., Sharma,R., Schoenberg,M.P. and Netto,G.J. (2012) High epidermal growth factor receptor immunohistochemical expression in urothelial carcinoma of the bladder is not associated with EGFR mutations in exons 19 and 21: a study using formalin-fixed, paraffin-embedded archival tissues. *Hum. Pathol.*, **43**, 1590–1595.
31. Røtterud,R., Nesland,J.M., Berner,A. and Fosså,S.D. (2005) Expression of the epidermal growth factor receptor family in normal and malignant urothelium. *BJU Int.*, **95**, 1344–1350.
32. Nedaeinia,R., Avan,A., Manian,M., Salehi,R. and Ghayour-Mobarhan,M. (2014) EGFR as a potential target for the treatment of pancreatic cancer: dilemma and controversies. *Curr. Drug Targets*, **15**, 1293–1301.
33. Harbeck,N., Penault-Llorca,F., Cortes,J., Gnant,M., Houssami,N., Poortmans,P., Ruddy,K., Tsang,J. and Cardoso,F. (2019) Breast cancer.
34. Han,J., Liu,Y., Yang,S., Wu,X., Li,H. and Wang,Q. (2021) MEK inhibitors for the treatment of non-small cell lung cancer. *J. Hematol. Oncol.*, **14**, 1.
35. Xie,Y.-H., Chen,Y.-X. and Fang,J.-Y. (2020) Comprehensive review of targeted therapy for colorectal cancer. *Signal Transduct. Target. Ther.*, **5**, 22.
36. Johnson,D.E., Burtness,B., Leemans,C.R., Lui,V.W.Y., Bauman,J.E. and Grandis,J.R. (2020) Head and neck squamous cell carcinoma. *Nat. Rev. Dis. Prim.*, **6**, 92.
37. Mirza-Aghazadeh-Attari,M., Ostadian,C., Saei,A.A., Mihanfar,A., Darband,S.G., Sadighparvar,S., Kaviani,M., Samadi Kafil,H., Yousefi,B. and Majidinia,M. (2019) DNA damage response and repair in ovarian cancer: Potential targets for therapeutic strategies. *DNA Repair (Amst)*., **80**, 59–84.
38. Al-Alem,L.F., Baker,A.T., Pandya,U.M., Eisenhauer,E.L. and Rueda,B.R. (2019) Understanding and Targeting Apoptotic Pathways in Ovarian Cancer. *Cancers* , **11**.
39. Zervantonakis,I.K., Iavarone,C., Chen,H.-Y., Selfors,L.M., Palakurthi,S., Liu,J.F., Drapkin,R., Matulonis,U., Leverson,J.D., Sampath,D., *et al.* (2017) Systems analysis of apoptotic priming in ovarian cancer identifies vulnerabilities and predictors of drug response. *Nat. Commun.*, **8**, 365.
40. Shen,Q., Li,J., Mai,J., Zhang,Z., Fisher,A., Wu,X., Li,Z., Ramirez,M.R., Chen,S. and Shen,H. (2018) Sensitizing non-small cell lung cancer to BCL-xL-targeted apoptosis. *Cell Death Dis.*, **9**, 986.
41. Han,Z., Liang,J., Li,Y. and He,J. (2019) Drugs and Clinical Approaches Targeting the Antiapoptotic Protein: A Review. *Biomed Res. Int.*, **2019**, 1212369.
42. Du,R., Huang,C., Liu,K., Li,X. and Dong,Z. (2021) Targeting AURKA in Cancer: molecular mechanisms and opportunities for Cancer therapy. *Mol. Cancer*, **20**, 15.
43. Umene,K., Yanokura,M., Banno,K., Irie,H., Adachi,M., Iida,M., Nakamura,K., Nogami,Y., Masuda,K., Kobayashi,Y., *et al.* (2015) Aurora kinase A has a significant role as a therapeutic target and clinical biomarker in endometrial cancer. *Int. J. Oncol.*, **46**, 1498–1506.
44. Do,T.-V., Xiao,F., Bickel,L.E., Klein-Szanto,A.J., Pathak,H.B., Hua,X., Howe,C., O'Brien,S.W., Maglaty,M., Ecsedy,J.A., *et al.* (2014) Aurora kinase A mediates epithelial ovarian cancer cell migration and adhesion. *Oncogene*, **33**, 539–549.
45. Pérez-Fidalgo,J.A., Gambardella,V., Pineda,B., Burgues,O., Piñero,O. and Cervantes,A. (2020) Aurora kinases in ovarian cancer. *ESMO Open*, **5**, e000718.
46. Shah,K.N., Bhatt,R., Rotow,J., Rohrberg,J., Olivas,V., Wang,V.E., Hemmati,G., Martins,M.M., Maynard,A., Kuhn,J., *et al.* (2019) Aurora kinase A drives the evolution of resistance to third-generation EGFR inhibitors in lung cancer. *Nat. Med.*, **25**, 111–118.
47. Takai,N., Desmond,J.C., Kumagai,T., Gui,D., Said,J.W., Whittaker,S., Miyakawa,I. and Koeffler,H.P. (2004) Histone Deacetylase Inhibitors Have a Profound Antigrowth Activity in Endometrial Cancer Cells. *Clin. Cancer Res.*, **10**, 1141–1149.
48. Psilopatis,I., Pergaris,A., Giaginis,C. and Theocharis,S. (2021) Histone Deacetylase Inhibitors: A Promising Therapeutic Alternative for Endometrial Carcinoma. *Dis. Markers*, **2021**, 7850688.
49. Moufarrij,S., Srivastava,A., Gomez,S., Hadley,M., Palmer,E., Austin,P.T., Chisholm,S., Diab,N., Roche,K., Yu,A., *et al.* (2020) Combining DNMT and HDAC6 inhibitors increases anti-tumor immune signaling and decreases tumor burden in ovarian cancer. *Sci. Rep.*, **10**, 3470.
50. Yang,H., Sun,B., Xu,K., He,Y., Zhang,T., Hall,S.R.R., Tan,S.T., Schmid,R.A., Peng,R.-W., Hu,G., *et al.* (2021) Pharmaco-transcriptomic correlation analysis reveals novel responsive signatures to HDAC inhibitors and identifies Dasatinib as a synergistic interactor in small-cell lung cancer. *eBioMedicine*, **69**.
51. Leonardi C.,G., Falzone,L., Salemi,R., Zanghì,A., Spandidos A.,D., Mccubrey A.,J., Candido,S. and Libra,M. (2018) Cutaneous melanoma: From pathogenesis to therapy (Review). *Int J Oncol*, **52**, 1071–1080.
52. Hodis,E., Watson,I.R., Kryukov,G.V., Arold,S.T., Imielinski,M., Theurillat,J.-P., Nickerson,E., Auclair,D., Li,L., Place,C., *et al.* (2012) A Landscape of Driver Mutations in Melanoma. *Cell*, **150**, 251–263.
53. Ghosh,A.P. and Sudarshan,S. (2016) Genetics of renal cancer: focus on MTOR. *Aging (Albany. NY).*, **8**, 421–422.
54. Miricescu,D., Balan Gabriela,D., Tulin,A., Stiru,O., Vacaroiu Adela,I., Mihai Andrada,D., Popa Constantin,C., Papacocea Ioana,R., Enyedi,M., Sorin Andrei,N., *et al.* (2021) PI3K/AKT/mTOR signalling pathway involvement in renal cell carcinoma pathogenesis (Review). *Exp Ther Med*, **21**, 540.
55. Qin,A., Musket,A., Musich,P.R., Schweitzer,J.B. and Xie,Q. (2021) Receptor tyrosine kinases as druggable targets in glioblastoma: Do signaling pathways matter? *Neuro-Oncology Adv.*, **3**, vdab133.
56. Alexandru,O., Horescu,C., Sevastre,A.-S., Cioc,C., Baloi,C., Oprita,A. and Dricu,A. (2020) Receptor tyrosine kinase targeting in glioblastoma: performance, limitations and future approaches. *Contemp. Oncol. Onkol.*, **24**, 55–66.
57. Ben-David,U., Siranosian,B., Ha,G., Tang,H., Oren,Y., Hinohara,K., Strathdee,C.A., Dempster,J., Lyons,N.J., Burns,R., *et al.* (2018) Genetic and transcriptional evolution alters cancer cell line drug response. *Nature*, **560**, 325–330.
58. Di Veroli,G.Y., Fornari,C., Wang,D., Mollard,S., Bramhall,J.L., Richards,F.M. and Jodrell,D.I. (2016) Combenefit: an interactive platform for the analysis and visualization of drug combinations. *Bioinformatics*, **32**, 2866–2868.
59. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**.
60. Korotkevich,G., Sukhov,V., Budin,N., Shpak,B., Artyomov,M.N. and Sergushichev,A. (2021) Fast gene set enrichment analysis. *bioRxiv*, 10.1101/060012.


A
CCLs
Genes
Gene Expression
ρ
Drugs
CCLs
Drug Potency (CTRP2)
Drug Potency (GDSC)
Drug Potency (PRISM)
Drug 1
Drug N
PCC
Genes
Genomic Profiles of Drug Sensitivity (GPDS)
Gene$_X$ Expression
Drug$_Y$ Potency
Marker of resistance
Marker of sensitivity
B
Cells
Genes
RAW UMI Counts
GF-ICF normalization
Cell Profile
GPDS of drug X
High
Low
GF-ICF value
GSEA
sensitive cell
tolerant cell
C
number of genes as GSEA input
50
100
250
500
1000
CTRP2
GDSC
PRISM
Precision
Recall
Sensitivity
1 − Specificity
··· random
D
Compounds
Cell lines
AUC of ROC curve
CTRP2
GDSC
PRISM
GDSC1
E
F


A
Cell-lines
C1
C2
C3
C4
Drug Dataset
D1
D2
D3
D4
Compounds
Cancer Type
Cancer Type
Lung
Skin
Breast
H&N
Colorectal
Pancreatic
Other
Drug Dataset
CTRP2
GDSC
PRISM
% of predicted sensitive cells
100
90
B
D1
EGFR
ERBB2
MAP2K1
AKT1
MAP2K2
NR3C1
AKT2
AKT3
PIK3CA
KDR
D2
BCL2
BCL2L1
HDAC1
KDR
AURKA
HDAC6
BCL2L2
AURKB
HDAC2
HDAC3
D3
HDAC1
MTOR
KIT
TUBB
BRAF
HDAC3
KDR
PDGFRB
FLT3
HDAC2
D4
KDR
KIT
FLT3
FLT1
GSK3B
MET
PIK3CD
BRAF
FGFR1
MTOR
number of drugs
Targeted genes
C
Bortezomib
Time (h)
0
12
48
96
UMAP 2
UMAP 1
D
density
0
12
48
96
sensitive
resistant
Bortezomib ES score
E
% of cells
Bortezomib
DREEP prediction
Resistant
Sensitive
t0
t12
t48
t96
F
Bortezomib
Time (h)
0
12
48
96
UMAP 2
UMAP 1
Bortezomib ES
sensitive
resistant
UMAP 1


A
Cell lineage
afatinib-tolerant
afatinib-sensitive
UMAP 2
UMAP 1
B
specific for afatinib-tolerant lineages
specific for afatinib-sensitive lineages
median ES (afatinib-sensitive lineages)
median ES (afatinib-tolerant lineages)
C
Median ES
tolerant
sensitive
Drugs
afatinib-sensitive lineages
afatinib-tolerant lineages
MoA
D
Syngergy Score (Median)
Tenovin−6 (SIRT1,2)
Uprosertib (AKT)
Belinostat (HDAC1)
PIK93 (PI3Kγ)
AR−42 (HDAC1)
Buparlisib (PI3Kα,β,γ,δ)
Mode of Action (MoA)
EGFR signaling
AKT/MTOR/PI3K inhibitors
HDAC inhibitors
IGF1R signaling
E
Cell lineage
afatinib-tolerant
afatinib-sensitive
UMAP 2
UMAP 1