Patient-Derived Xenografts of Triple-Negative Breast Cancer Enable Deconvolution and Prediction of Chemotherapy Responses

Jonathan T. Lei[1,2], Lacey E. Dobrolecki[1], Chen Huang[1,15], Ramakrishnan R. Srinivasan[1], Suhas V. Vasaikar[1,16], Alaina N. Lewis[1], Christina Sallas[1], Na Zhao[2], Jin Cao[1,18], John D. Landua[1], Chang In Moon[1,3], Yuxing Liao[1,3], Susan G. Hilsenbeck[1], C. Kent Osborne[1], Mothaffar F. Rimawi[1], Matthew J. Ellis[1,18], Varduhi Petrosyan[3], Bo Wen[1,19], Kai Li[1,20], Alexander B. Saltzman[1,4], Antrix Jain[1,4], Anna Malovannaya[1,4,5], Gerburg M. Wulf[6], Elisabetta Marangoni[7], Shunqiang Li[8], Daniel C. Kraushaar[2], Tao Wang[1], Senthil Damodaran[9], Xiaofeng Zheng[9], Funda Meric-Bernstam[9], Gloria V. Echeverria[1,2,10,11], Meenakshi Anurag[1,10], Xi Chen[1,9], Bryan E. Welm[12], Alana L. Welm[13], Bing Zhang[1,3,*], Michael T. Lewis[1,2,14,21,*]

[1]Lester and Sue Smith Breast Center and Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA

[2]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA

[3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA

[4]Mass Spectrometry Proteomics Core, Baylor College of Medicine, Houston, TX 77030, USA

[5]Department of Biochemistry and Molecular Pharmacology, Baylor College of Medicine, Houston, TX 77030, USA

[6]Cancer Research Institute, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA

[7]Laboratory of Preclinical investigation, Translational Research Department, Institut Curie, PSL University, 26 Rue d'Ulm, Paris 75005, France

[8]Siteman Cancer Center, Department of Medicine, Washington University in St. Louis, St. Louis, MO 63108, USA

[9]MD Anderson Cancer Center, Houston, TX 77030, USA

[10]Department of Medicine, Baylor College of Medicine, Houston, TX 77030, USA

[11]Department of Radiation Oncology, Baylor College of Medicine, Houston, TX 77030, USA

[12]Department of Surgery, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA

[13]Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA

[14]Department of Radiology, Baylor College of Medicine, Houston, TX 77030, USA

[15]Current affiliation: Department of Genetics, University of Alabama at Birmingham, Birmingham, AL 35294, USA

[16]Current affiliation: Translational Oncology Bioinformatics, Pfizer, Bothell, WA 98021, USA

[17]Current affiliation: The MOE Key Laboratory of Biosystems Homeostasis & Protection and Innovation Center for Cell Signaling Network, Life Sciences Institute, Zhejiang University, Hangzhou 310058, China

[18]Current affiliation: Guardant Health, Palo Alto, CA 94304, USA

[19]Current affiliation: Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

[20]Current affiliation: Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA

[21]Lead contact

*Correspondence: mtlewis@bcm.edu (M.T.L.) and bing.zhang@bcm.edu (B.Z.)

Summary

Combination chemotherapy remains essential for clinical management of triple-negative breast cancer (TNBC). Consequently, responses to individual agents cannot be easily delineated at the single patient level, even though some patients might not require all drugs in the combination. Herein, we conduct multi-omic analyses of orthotopic TNBC patient-derived xenografts (PDXs) treated with single agent carboplatin, docetaxel, or the combination. Combination responses were usually no better than the best single agent, with enhanced response in only ~13% of PDX, and apparent antagonism in a comparable percentage. Single-omic comparisons showed largely non-overlapping results between genes associated with single agent and combination treatments that could be validated in independent patient cohorts. Multi-omic analyses of PDXs identified agent-specific biomarkers/biomarker combinations, nominating high Cytokeratin-5 (KRT5) as a general marker of responsiveness. Notably, integrating proteomic with transcriptomic data improved predictive modeling of pathologic complete response to combination chemotherapy. PDXs refractory to all treatments were enriched for signatures of dysregulated mitochondrial function. Targeting this process indirectly in a PDX with HDAC inhibition plus chemotherapy *in vivo* overcomes chemoresistance. These results suggest possible resistance mechanisms and therapeutic strategies in TNBC to overcome chemoresistance, and potentially allow optimization of chemotherapeutic regimens.

Highlights

1. Minable multi-omic resource of baseline TNBC PDX tumors matched with single and combination chemotherapy responses yields candidate resistance mechanisms and potentially targetable processes
2. Combination carboplatin/docetaxel is largely ineffective at enhancing response versus the best single agent in TNBC PDXs, and they can antagonize one another in some PDXs
3. Proteomics data enhances predictions of platinum- and taxane-based chemotherapy response
4. Targeted agents can enhance chemotherapy response in TNBC PDXs

Introduction

Breast cancer is a leading cause of cancer-related deaths worldwide. Triple-negative breast cancer (TNBC) is characterized by the absence of estrogen receptor (ER), progesterone receptor (PR), and low/no expression of HER2. Consequently, there is a lack of effective molecularly targeted therapeutics and chemotherapy remains the mainstay of systemic therapy. In the neoadjuvant setting, taxanes (paclitaxel, docetaxel), anthracyclines (doxorubicin, epirubicin), cyclophosphamide, and platinums (cisplatin, carboplatin) are used, often in series and/or in combination. Identifying patients who will respond to chemotherapy, and finding optimal chemotherapy regimens to treat TNBC, is an area of intense investigation. Recently, approved immune checkpoint blockade inhibitors are now available for TNBC patients and PARP inhibition has also been approved as adjuvant therapy for the ~15% of TNBC patients harboring deleterious BRCA1/2 mutations[1].

The addition of taxane to adriamycin/cyclophosphamide (AC)-based regimens improved outcomes[2,3], and was the standard-of-care to treat breast tumors for many years. The discovery that a subset of TNBC may be sensitive to DNA damaging agents[4] prompted the investigation of platinum derivatives for TNBC. Most recently, taxane/platinum combinations with or without AC have proven promising as they improve pathologic responses above historic single-agent rates. Although the exact treatment regimen was different among clinical trials such as GeparSixto[5], BrighTNess[6], and CALGB 40603[7], TNBC patients enrolled in treatment arms receiving combination platinum and taxane achieved pCR rates of over 50%, which was significantly higher than the pCR rate of patients in arms receiving either drug as a single agent.

Because platinum-based agents can increase the frequency of adverse events[8], more current studies such as NeoCART[9], and two TNBC clinical trials (NCT02547987 and NCT02124902), analyzed by the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC), have investigated less intensive neoadjuvant regimens for TNBC patients consisting of platinum + taxane and excluding anthracyclines[10]. However, in ~50% of cases, pCR is not achieved, and these individuals experience dramatically shorter relapse-free and overall survival[11]. Furthermore, it is unclear whether the two drugs interact to enhance responses beyond an additive effect. Results from the KEYNOTE-522[12] TNBC trial led to approval of adding an immune checkpoint inhibitor, pembrolizumab, to an already intensive chemotherapy backbone of carboplatin + paclitaxel followed by AC. However, addition of pembrolizumab improved the pCR rate by only ~7% compared to the control arm[13]. Thus, identification of predictive markers for the efficacy of individual drugs would be useful in order to enhance response rates and to optimize and individualize treatments, as all chemotherapy agents have potentially life-threatening short- and long-term toxicities.

Our previous study identified weighted gene co-expression network analysis followed by Connect The Dots (WGCNA/CTD) informed mRNA-based biomarker panels that can predict differential response to single agent platinums and taxanes in multiple PDX and clinical datasets[14]. However, these predictive models were evaluated by leave-one-out cross-validation (LOOCV) without performance assessment of a fixed model on unseen,

independent test data. Predictive biomarkers for the platinum + taxane combination were not identifiable. Furthermore, the mRNA-based biomarker panel did not provide insight into potential resistance mechanisms or therapeutic targets. Here, we analyzed treatment response data to single agent carboplatin and docetaxel for a cohort of 50 TNBC PDXs derived from 46 different TNBC patients. Of these, we also generated combination response data for 42 PDXs to carboplatin + docetaxel combination. Treatment response data were integrated with baseline proteogenomic (DNA/mRNA/protein) profiles to perform a set of unbiased analyses to address several key questions: 1) What are the molecular associates of and predictors of response and resistance to single agent and combination treatments? 2) What molecular features can inform the selection of a single agent for individual tumors? 3) For the few PDXs where combination was beneficial vs single agents, what features are associated with these tumors, and what features are associated with cases where the combination was worse than the best single agent, and 4) What are possible therapeutic options for tumors that lack treatment response to all chemotherapy regimens?

Results

Combination carboplatin and docetaxel is largely ineffective at generating enhanced responses over the best single agent in TNBC PDXs

TNBC patients frequently receive combination chemotherapy treatments based on clinical data suggesting improved response from combination chemotherapy compared with single-agent chemotherapy[2,3]. However, averaging effects observed in group level data from clinical trials may not be applicable to individual patients[15]. Further, in clinical investigations, it is impossible to assign one patient to more than one treatment arm to deconvolute response to individual agents versus combinations. We hypothesized that this limitation can potentially be addressed through preclinical trials using PDX models.

To test this hypothesis, and to exploit this unique opportunity, we leveraged our collection of data from 49 TNBC PDX models and one estrogen-independent, low estrogen receptor-expressing model (FCP699) (reported to us originally as a TNBC, and propagated without estrogen as such until ER immunostaining showed low levels of ER in some nuclei) aggregated from three separate preclinical trials (**Figure S1A** and **Table S1A-B**). These trials included treatment arms consisting of control (no treatment) versus four weekly cycles of human-equivalent doses of carboplatin (50 mg/kg), docetaxel (20 mg/kg) or the combination at these doses[14]. Treatment effects were evaluated by $\log_2$ fold-change in tumor volume after four weeks vs. baseline and compared using general linear models (**Methods** and **Table S1B-D**). Quantitative and qualitative best clinical responses (**Methods**) to each treatment were summarized for each PDX model (**Figure 1A**) and plotted as individual tumor volumes over time (**Figure S1B**). All PDXs had treatment responses documented for single agent carboplatin and docetaxel, while 84% (42/50) of PDXs had combination response information.

Among PDXs that had treatment response information for all arms (**Figure 1A**, Q1), 71% (30/42) were responsive (complete response [CR] or partially responsive [PR]) to any treatment arm, while 29% (12/42) were resistant (stable disease [SD] or progressive disease [PD]) to all treatment arms, according to our modified RECIST criteria (mRECIST)[16,17]. When including the full cohort of 50 PDXs, 76% (38/50) were responsive to any treatment. When comparing quantitative tumor responses based on tumor volume change, single agent docetaxel led to stronger tumor shrinkage in a greater number of PDX models compared to carboplatin at the doses evaluated (n = 18 vs n = 11, respectively, **Figure 1A**, Q2).

As expected, overall average tumor shrinkage was significantly greater in the combination arm compared to either the carboplatin (Wilcoxon signed rank test $p = 3.19e{-}06$) or docetaxel arms (Wilcoxon signed rank test $p = 0.02$) (**Figure 1B**). However, average shrinkage in the combination arm was generally comparable to the best single agent response for each PDX ($p = 0.70$) (**Figure 1B**). This result was confirmed orthogonally with a Bliss independence model[18] which showed no significant difference between expected and observed combination tumor responses (paired t-test $p = 0.12$ and Wilcoxon signed rank test $p = 0.42$) (**Figure S1C** and **Methods**), further supporting the observation that the increased efficacy of combination comes from an additive, not

synergistic, benefit. Individually, combination treatment generated significantly enhanced responses vs the best single agent in only 13% (4/30, $p < 0.05$) of PDXs (**Figure 1A,** Q3). Interestingly, antagonism was also observed for combination treatment where combination response was significantly worse ($p < 0.05$) compared to the best single agent response in 12% (5/42) of PDXs (**Figure 1A,** Q3). Moreover, in 97% (29/30) of PDXs that had a response to any of the three arms tested, combination treatment did not qualitatively improve response assessed by mRECIST compared to the best single agent treatment (**Figure 1A**). These results strongly suggest that the standard-of-care for TNBC patients consisting of multiple chemotherapy agents can be optimized to reduce ineffective (unnecessary) treatment, as well as unwanted side effects, if patients could be matched to the best single agent treatment accurately.

Molecular associates of chemotherapy response

To identify molecular correlates of treatment response and resistance that might provide mechanistic insight, we examined associations between mutation, copy number, mRNA, and protein profiles in baseline PDX tumors (**Table S2**) with treatment responses (**Table S3A-S3O**). Molecular profiling by whole exome sequencing, RNA-Seq, and mass spectrometry-based proteomic profiling in these PDXs has been previously described[14] and in **Methods**. Significant associations were observed for 2 non-silent mutations, 429 mRNAs, and 81 proteins for carboplatin response; 5 non-silent mutations, 222 mRNAs, and 44 proteins for docetaxel response; and 8 non-silent mutations, 190 mRNAs, and 55 proteins for combination treatment response ($p < 0.01$, Methods, **Figure 2A**, **Figure S2,** and **Table S3P**). No significant associations were found with copy number data ($p < 0.01$). At a more relaxed $p < 0.05$ threshold, high-level amplification events in 39 genes associated with carboplatin response, 22 genes associated with docetaxel resistance, and none were associated with combination response (**Figure S2**). Many genes found to be associated with carboplatin response were previously implicated in platinum resistance in cancer[19] (**Table S3Q**). Strikingly, molecular associates of response to individual treatment regimens had only limited overlap for individual data types (**Figure 2A**). This was also the case when considering all omic platforms together (**Figure 2B** and **Table S3R**), consistent with distinct molecular mechanisms underlying response to each treatment. As an example, deleterious *BRCA2* variants were associated with increased carboplatin response (**Figure S2A**) but not to docetaxel, nor to the combination. Deleterious *BRCA1* variants also showed this trend but were not significant in this analysis (**Figure S2A**). There were some associations that were shared between different treatments. For example, Glutamate Rich 6 (ERICH6) variants were associated with increased response to both single-agent carboplatin and docetaxel, but not the combination (**Figure 2B**, **Figure S2A-C**). For the same treatment, different types of molecular data identified distinct sets of associated genes (**Figure 2C, Table S3S**), each contributing unique yet complementary information. This finding is not unexpected given that different types of molecular data are not fully correlated with one another. For example, the median gene-wise Spearman's correlation between mRNA and protein was 0.46, highlighting the poor mRNA-protein correlation observed for many genes in this and other proteogenomic studies[20]. This underscores the importance of multi-omic analyses to gain a comprehensive understanding on the genes that may underlie treatment response.

To examine potential clinical relevance of genes associated with different treatment regimens at either mRNA or protein levels in our PDX analysis, we next annotated genes identified in our study using data from independent PDX and clinical trial datasets for TNBC samples that received similar treatments before response assessment (**Table S4A**). We prioritized genes that were found to be associated in at least 3 or more datasets, including at least one of our PDX-based datasets (**Figure 2D-F**). Since concordant associations across multiple datasets reinforce each other, we relaxed p-value thresholds from $p < 0.01$ to $p < 0.05$. We also annotated these genes by our drug target tier system that may further nominate candidates for experimental followup or therapeutic development[20]. In addition, we annotated "functionally dark" genes that represent understudied (or unstudied) genes based on publication count[21].

Platinum treatment datasets included two PDX datasets with cisplatin (GSE142767[22] and Curie Institute [unpublished]), a cisplatin treatment arm from a clinical trial (GSE18864[23] [DFCI]), and manually curated genes from the literature that have been associated with platinum resistance (Huang et al, 2021[19]) filtered to those with an association in xenograft or in patient specimens (**Figure 2D**). Among the genes consistently higher in platinum-refractory samples was Glutathione Synthetase (GSS), a target of an approved non-oncology drug. The impact of non-oncology drugs on cancer viability has recently come into focus[24] and further investigation of how perturbation of GSS may augment platinum response could drive future drug repurposing efforts. Lower level of Glucoside Xylosyltransferase 1 (GXYLT1), a "dark" gene, was consistently associated with platinum responsiveness. GXYLT1 is understudied in cancer and could potentially be a candidate for further study to better understand markers of tumor response to chemotherapy and to platinum agents in particular.

Taxane treatment datasets included a study with paclitaxel treated PDX tumors (GSE142767[22]) a paclitaxel containing arm from the I-SPY2 trial[25] (GSE194040) with tumor responses recorded after paclitaxel only before AC treatment[26], and from a paclitaxel arm from a previously completed clinical trial[27] in which responses were also recorded before AC treatment (**Figure 2E**). Another "functionally dark" gene, C16orf72 was found to be associated with taxane resistance which suggests another understudied gene for future studies to expand our understanding of taxane resistance genes.

Combination datasets were from a CPTAC study where TNBC patients were treated with carboplatin + docetaxel[10], and from the City of Hope (COH) National Medical Center where patients were treated with carboplatin + paclitaxel[28] (**Figure 2F**). We observed an understudied gene, FHF Complex Subunit HOOK Interacting Protein 2A (FAM160B1), that might have previously unknown roles in combination chemotherapy responses in TNBC. Among the genes consistently associated with chemotherapy response, only E2F Transcription Factor 7 (E2F7) was associated with more than one treatment type (**Figure 2G**). This is consistent with our above observation that distinct genes may regulate tumor responsiveness and resistance to different chemotherapy treatments. At the pathway level, RNA and protein data converged on upregulated metabolic pathways in PDXs resistant to all treatment types while upregulated E2F/G2M related pathways associated with PDXs responsive to all treatment types. Variable levels of MYC-related proliferation and interferon pathways

were observed across different treatments, suggesting pathways that are common across, and unique to, different treatments (**Figure 2H**).

Predicting chemotherapy response

Based on the observation that a limited number of gene sets were associated with response (**Figure 2H)**, machine learning models were employed to predict pathologic complete response (pCR) to chemotherapy using the Protein Marker Selection (ProMS) tool we recently developed[29]. ProMS identifies a group of minimally redundant, and therefore complementary, markers associated with an attribute label, which for this study is chemotherapy response, where each marker represents a biological function or pathway defined by co-expressed genes. Although initially developed for protein marker selection using single or multi-omics data, ProMS can also be implemented for RNA marker selection, which is how it was applied in this study. Individual datasets from this study, and previously described external resources (**Table S4B**), were used to train separate logistic regression models using ProMS selected features to predict CR/pCR for platinum, taxane, and their combination, respectively, using transcriptomics profiling data (**Figure S3A**). Due to the small sample size of each dataset, five features were selected for model training. For each treatment type, a clinical dataset, or the largest clinical dataset when more than one exists, was selected as the independent test data to evaluate fully trained model performance. Prediction performance of models trained using single datasets were highly variable and close to random (0.50) as measured by area under ROC curve (AUROC) in the set-aside cross-validation data from the same dataset (**Figure S3B**). In particular, the taxane dataset 06 from Savage et al, 2020[22] had a low AUROC of 0.35 and was omitted in subsequent analyses.

Since models trained on individual datasets were unstable and not predictive, likely due to interpatient heterogeneity in chemotherapy resistance pathways, which cannot be fully captured by small sample sizes, datasets of the same treatment type were batch corrected and merged to increase sample size (**Figures S3C-F**). Trained logistic regression models using up to 10 ProMS selected features (10% of training sample size) on the merged platinum training data, showed the best AUROC in the set-aside cross-validation merged data with 10 markers (**Figure 3A** and **Table S4C**). The fully trained platinum predictor achieved an AUROC of 0.59 on the independent test dataset 04 from Silver et al, 2010[23] (**Figure 3C**). A similar procedure was performed for the taxane (**Figures 3D-F**) and platinum + taxane (**Figures 3G-I**) predictors, where logistic regression models with 6 and 10 ProMS selected markers were found achieve the best cross-validated performance, respectively (**Figure 3D** and **Figure 3G**). Performance of fully trained taxane and platinum + taxane logistic regression models with ProMS selected features achieved an AUROC of 0.67 (**Figure 3F**) and 0.74 (**Figure 3I**) on independent test data, respectively.

The ProMS method is unique for its ability to be extended to the multi-omics setting, leveraging multi-omic data for biomarker discovery within a single omic layer of interest[29]. For the platinum + taxane predictor, both RNA and protein data were available for training datasets. Since 10 biomarkers achieved the best cross-validated performance by using RNA data alone (**Figure 3G**), merged RNA and protein data (**Figure S3E-F**) were used

together to also select a group of 10 protein-informed, RNA biomarkers by running ProMS in multi-omics mode ($ProMS^{mo}$). Interestingly, only one of these genes, LGALS1, overlapped with the RNA-only informed ProMS model (**Figure 3H** and **Table S4C**). Performance of the platinum + taxane predictor trained with features selected by $ProMS^{mo}$ outperformed ProMS on independent test data, achieving an AUROC of 0.85 (**Figure 3I**). These results suggest protein data heavily impacts RNA marker selection and can enhance prediction performance when used together with RNA data. It is notable that the protein-informed model achieved a 0% false positive rate with close to 60% sensitivity, as revealed by the sharp initial climb in the AUROC curve, underscoring its high specificity in predicting pCR tumors in response to platinum + taxane chemotherapy.

To generate predictive models that can be taken forward in future PDX-based studies and clinical trials, final, composite, logistic regression models for prediction of platinum, taxane, and platinum + taxane response were generated using all existing data. This contrasts with the models described above where a clinical dataset was withheld intentionally to use as an independent test dataset. These composite models represent "the best that can be done" until additional datasets become available. Logistic regression models were trained using ProMS selected features from **Figures 3B, 3E,** and **3H**, respectively with all existing data. Average AUROC performance in the set-aside cross-validation for the composite predictors were 0.77, 0.68, and 0.78 for the platinum, taxane, and combination predictors, respectively (**Figure 3J**). All composite models outperformed respective non-composite models from **Figures 3A**, **3D**, and **3G**. Composite models were also less variable and therefore more stable across cross-validation repeat testing, especially for the taxane and combination models. The performance of these models was examined with additional independent datasets generated by our group. Using RNA-Seq data and chemotherapy response for eight additional TNBC PDXs independent of the 50 reported in this study (**Figure S3G-H**), the composite single agent platinum and taxane predictors achieved an AUROC performance of 0.57 and 0.71, respectively (**Figure 3K-L**). In addition, we analyzed RNA-Seq data from previously collected baseline TNBC tumors with neoadjuvant responses to single agent docetaxel[30] where the taxane predictor achieved an AUROC of 0.83 when applied to this dataset (**Figure 3L** and **Figure S3H**). Responsive models were very limited in these datasets which likely contributed towards variable performance. Robust performance of these composite predictors await further evaluation in datasets that may become available in the near future, including data from the BEAUTY[31], TBCRC 030[32], and RESPONSE (NCT05020860) clinical trials, or in future experimental settings using PDX/PDX-derived organoids.

In sum, these results highlight the potential of using machine learning models, and the integration of protein data when available, to select TNBC tumors that may respond to either single agent carboplatin, docetaxel or their combination, consistent with our previous work[14].

Molecular features that discriminate best single agent response to either carboplatin or docetaxel

To identify molecular features that discriminate single agent response to carboplatin or docetaxel directly, we identified 11 models for which response to carboplatin was significantly better than docetaxel, and 18 models in which response to docetaxel was significantly better than carboplatin (**Figure 1A** [Q2] and **Table S1C**).

Comparisons between the 11 and 18 models were performed using mutation, copy number, RNA, and protein data at the gene (**Figure S4A** and **Table S5A-D**) and pathway (**Figure 4A** and **Table S5E-J**) levels. In PDXs where carboplatin response was better than docetaxel, molecular data, particularly RNA and protein, was associated with up-regulation of gene sets broadly related to transcriptional regulation and cell cycle, among others (**Figure 4A**). Gene level changes reinforced geneset findings where individual genes were elevated in PDXs whose response to carboplatin was better than docetaxel such as Zinc Finger Protein 367 (ZNF367) (transcriptional regulation gene), and Anaphase Promoting Complex Subunit 5 (ANAPC5) (cell cycle gene) (**Figure S4B**). These genes at the RNA level were also associated individually with neoadjuvant cisplatin response in TNBC tumors from the DFCI cohort where the AUROC was the highest compared to AUROCs for the same genes in other neoadjuvant TNBC clinical datasets where patients received either single agent taxane or carboplatin + taxane before response assessment (**Figure 4B**). In contrast, in PDXs for which docetaxel response was better than carboplatin, both RNA and protein data associated higher levels of genes positively regulating the microtubule and actin cytoskeleton such as Rac Family Small GTPase 1 (RAC1), Peptidylprolyl Isomerase B (PPIB), Migration And Invasion Enhancer 1 (MIEN1), Coronin 1B (CORO1B), and Cofilin 1 (CFL1) (**Figure S4B**), suggesting microtubule-actin coordination[33] may impact responses to taxanes. In addition, elevated levels of GSS were also observed in PDXs with better response to docetaxel vs carboplatin (**Figure S4B**), corroborating earlier results that associated GSS with carboplatin resistance in multiple datasets (**Figure 2D**). Several genes at the RNA level were also found to be most discriminatory in the MDA patient dataset where response was assessed after receiving single-agent paclitaxel neoadjuvant chemotherapy (**Figure 4B**). A similar observation was made for genes found to be associated at the RNA level in addition to one or more omic profiling platforms (**Figure S4C**). Discrepancies between the two taxane datasets may be due to differences in transcriptomic platforms as the MDA dataset was profiled by RNA-Seq while the ISPY2 was profiled by microarray. In sum, these results prioritize genes and pathways associated with response to individual agents with clinical relevance that can be further investigated to better understand the underlying biology behind different classes of chemotherapy treatment.

Actin dynamics, protein trafficking, and mitochondrial function are correlated with positive and negative interactions between single agents

In addition to the few models that showed benefit from combination treatment, there were also a comparable number of PDXs where the combination response was worse than that to the best single agent (12% [5/42]) (**Figure 1A**), a phenomenon observed previously in *in vitro* studies[34]. To examine molecular features associated with positive vs negative interaction between single agents, we took two statistical approaches. Kruskall-Wallis testing using RNA or protein measurements for genes across four groups (**Figure 1A**, Q3, **Table S1D**, and **Figure 5A**): 1) PDXs for which the combination was better than single agents (n = 4), 2) PDXs for which the combination was worse than the best single agent (n = 5), 3) PDXs in which there was no enhanced response to the combination compared to the best single agent (n = 21), and 4) PDXs which were resistant to all treatments (n = 12). Based on results from RNA and protein analyses, we further computed meta-p-values for each gene by combining p-values from transcriptomic and proteomic analyses and corrected for multiple comparisons using

the meta-p-values. Three genes (adjusted meta-p-values < 0.05) were associated by RNA and protein data with PDXs where the combination generated enhanced response vs single agents (**Figure 5B**). Coactosin Like F-Actin Binding Protein 1 (COTL1) and ELKS/RAB6-interacting/CAST Family Member (ERC1) were higher while Dynein Light Chain LC8-Type 2 (DYNLL2) was lower in PDXs where combination was beneficial vs all other groups. COTL1 has been shown to be a tumor suppressor in breast cancer by inhibiting non-canonical TGF-beta signaling[35] while ERC1 is involved in trafficking, and DYNLL2 has been reported to regulate retrograde movement of cargo along microtubules[36]. Three genes associated with PDXs where the combination was worse than the best single agent, including ATPase Plasma Membrane Ca^{2+} Transporting 4 (ATP2B4), ETHE1 Persulfide Dioxygenase (ETHE1), and Fumarylacetoacetate Hydrolase (FAH) (**Figure 5C**). ATP2B4 regulates calcium homeostasis, ETHE1 modulates hydrogen sulfide homeostasis in mitochondria, and FAH is involved in amino acid catabolism. Overall, these genes have roles in modulating actin dynamics, trafficking, and metabolism. Although understudied for their role in modulating combination chemotherapy responses in triple-negative breast cancer, our results prioritize these genes for further study as regulators of therapeutic efficacy.

In addition to multi-group comparisons, we also examined group comparisons for PDXs where the combination was beneficial were grouped vs all else and likewise examined group comparisons for PDXs where combination was worse than the best single agent were grouped vs all else (**Figure S5A-C** and **Table S6A-D**). RNA and protein data was also used for these group comparisons and then ranked by meta-p-values for each gene (**Figure S5D-E** and **Table S6E-F**). Genes associated from the multi-group comparison analysis were also top-associated genes from this analysis, reinforcing the reliability of these associations.

Using meta-p-values of all genes from the analysis from group comparisons **(Figure S5D-E** and **Table S6E-F)**, pathway enrichment analysis showed multiple mitochondrial oxidative phosphorylation (OXPHOS)-related gene sets were downregulated in PDXs where combo was beneficial (**Figure 5D** and **Table S6G-I**). This is consistent with the finding that oxidative phosphorylation along with fatty acid metabolism were among the top significantly enriched pathways associated with carboplatin + docetaxel resistance in the CPTAC-TNBC cohort[10]. In contrast, multiple gene sets related to chromatin regulation were up-regulated in PDXs where combo was worse than the best single agent (**Figure 5E** and **Table S6J-L**). These results suggest that such processes may interfere with optimal efficacy of combination treatment. We also found decreased MAPK signaling as well as decreased mitochondrial transcription and translation pathways in these PDXs (**Figure 5E**). MAPK signaling is well-established for platinum- and taxane-induced apoptosis, as well as for decreased mitochondrial transcription and translation, consistent with a previous study where mitochondrial protein synthesis defects inhibited DNA-damaging agent chemotherapy-induced apoptosis due to lack of respiratory complex proteins[37]. These results await further validation but may suggest these processes may underlie both the positive and negative interactions between carboplatin and docetaxel.

Molecular correlates to lack of response to any chemotherapy regimens evaluated

In PDXs that were treated with all 3 treatment arms, at least one treatment induced tumor shrinkage for most models, yet 29% (12/42) of PDXs were resistant to all chemotherapy regimens tested (**Figure 1A**). To identify molecular associations with lack of any chemotherapy response vs response to any treatment, we compared PDXs without a response to all treatment arms vs PDXs with tumor shrinkage due to any regimen using data generated across all omic platforms (**Figure S6A** and **Table S7A-D**). In PDXs resistant to all treatments, RNA-Seq and protein data associated genesets broadly related cellular respiration and metabolism, including mitochondrial transcription and translation, OXPHOS, and fatty acid and amino acid metabolism to resistance (**Figure 6A** and **Table S7E-J**). Of note, multiple proteostatic pathways related to protein folding, trafficking, and degradation along with unfolded protein response (UPR) pathways related to endoplasmic reticulum stress were heightened in PDXs resistant to all chemotherapy arms while chromatin regulation was downregulated in these PDXs. Both RNA and protein data supported various metabolism genes such as Ganglioside GM2 Activator (GM2A) involved in lipid metabolism, Galactose Mutarotase (GALM) involved in carbohydrate metabolism, and Carnitine O-Acetyltransferase (CRAT) involved in fatty acid homeostasis (**Figure 6B**). Upregulation of CRAT was confirmed to be associated with tumors resistant to neoadjuvant carboplatin + docetaxel in the CPTAC-TNBC clinical trial[10] where OXPHOS and fatty acid metabolism pathways were also found to be upregulated.

KRT5 is a chemotherapy response marker for carboplatin, docetaxel, and their combination

Interestingly, in PDXs responsive to any treatment arm, basal marker proteins and structural components of cells, such as cytokeratins KRT5, KRT6B, and KRT17 were among the most significantly upregulated genes especially at the protein level (**Figure 6C** and **Figure S6A**). These gene level results support the finding that cytoskeletal/ECM gene sets were found to be lower in PDXs resistant to all treatment arms (and therefore higher in PDXs responsive to any treatment arm) (**Figure 6A**). Since RNA and protein expression for these three markers was not significant ($p > 0.05$) using a traditional mean/median Wilcoxon test in human tumors from the CPTAC-TNBC clinical trial[10] **(Figure S6B**), we performed the Anderson-Darling (AD) test in non-pCR vs pCR tumors to focus on gene levels in tail regions of the distribution that may account for differences in a subset of samples. Of these 3 markers, the AD-test showed that only KRT5 was significantly different at the RNA level with protein data nearing $p < 0.05$ (**Methods**), a compatible result with our PDX data that suggests heterogeneity in KRT5 abundance for non-pCR tumors and that a subset of these non-pCR tumors may have lower levels of KRT5 compared to pCR tumors **(Figure S6C**). Both KRT5 mRNA and protein abundance were able to discriminate non-pCR CPTAC-TNBC samples from pCR samples with an AUROC of 0.69 and 0.65, respectively (**Figure S6D**). Importantly, both AUROC curves demonstrated appreciable sensitivity with a 0% false positive rate, indicating high specificity of KRT5 to distinguish non-pCR tumors to neoadjuvant chemotherapy. We further validated KRT5 by IHC staining of baseline tumors prior to treatment from our PDX cohort (**Figure 6D**) where staining intensity and frequency were quantified to generate an Allred score (**Figure 6D-F**). By using the Allred score, we achieved an AUROC of 0.83 in discriminating PDX tumors responsive to any chemotherapy arm in our cohort (**Figure 6F**) suggesting the potential use of KRT5 as a biomarker of responsiveness to non-anthracycline-based chemotherapy regimens consisting of carboplatin, docetaxel, or the combination.

Targeted agents can enhance chemotherapy response

In resistant PDXs, we also observed higher expression of several "cancer stem-like cell" associated pathways previously implicated in breast cancer, including Hedgehog, VEGF, Notch, and MET/RAP1/RAC1 (**Figure 6A**), inhibitors of which have all been evaluated in clinical trials, and all of which have not been approved for clinical use as single agents. A gamma secretase (Notch) inhibitor MK0752 was evaluated in combination with docetaxel and showed enhanced response but is not approved for clinical use[38]. For a partial analysis of these results we used our previously completed SPORE/U54 PDXNet preclinical trial (**Figure S1A**) in which we investigated responses to seven agents targeting these "cancer stem-like cell"-associated pathways alone and in combination with carboplatin in 11-20 randomly selected PDX models in a manner similar to unselected clinical trials without molecular matching (**Figure 7A**). Four models were included that were resistant to all chemotherapy regimens, BCM-0046, BCM-7821, BCM-15029, and HCI-028. Of the seven targeted agents evaluated, only niraparib (PARP inhibitor) and EPZ011989 (EZH2 inhibitor) resulted in significant tumor shrinkage (PR or better) as single agents, but in only two and one PDX models, respectively (**Figure 7A**). Of these three models responsive to single agents, only BCM-15029 was in the chemoresistant group. The response to single agents indicates a 2.6% overall success rate for candidate investigational agents in the absence of appropriate biomarkers. Although these sparse positive data prevent us from learning the biological mechanisms underlying drug response, we hypothesized that shared features between the responsive PDX models with other breast tumors that have high similarity to these PDXs could predict the observed therapeutic vulnerabilities.

To test this hypothesis, we applied a multi-omic computational approach that uses publicly available human data to prioritize agents for PDX testing by first querying the RNA profiles of the three PDX models that showed significant tumor shrinkage against the CPTAC cohort with proteogenomic profiling on breast tumors[39] (**Methods**). For each PDX, the top 20% of the CPTAC breast tumors with the highest similarity by RNA profiles to the PDX were considered as a PDX-like group. Since the direct targets of these agents are proteins, proteomics profiles were then compared to those from the remaining tumors to identify significantly up-regulated proteins as putative targets. Remarkably, when HCI-015, which showed a high sensitivity to Niraparib, was used as a query, PARP1 (target of niraparib) ranked first among all 841 quantified druggable proteins (**Figure 7B**). Moreover, for all three PDX models, the targets of effective agents (PARP1/niraparib for HCI-015 and BCM-15029, EZH2/EPZ011989 for BCM-15006) ranked the best among targets of all tested agents (**Figure 7B**) which was reflected in drug response profiles for these models (**Figure 7C**), providing strong support to our prediction approach. These data further suggest a single sample, tailored therapeutic strategy to treat a chemotherapy refractory TNBC PDX model, BCM-15029, and validates the general approach from our previous studies that selects PDX models for pharmacologic intervention based on drug target abundance[40–43].

Targeted agents were able to enhance carboplatin responses, albeit infrequently (**Figure 7A**). For example, consistent with single agent response to niraparib in BCM-15029, niraparib enhanced response to carboplatin in this chemoresistant model (**Figure 7A**). In another chemotherapy refractory model, BCM-0046, both romidepsin, an HDAC inhibitor, and carboplatin were completely ineffective as single agents, however, the combination

induced a partial response (**Figure 7D**). Protein abundance of romidepsin's targets HDAC1, HDAC2, and HDAC3 were highest in BCM-0046 out of all PDXs (**Figure 7E**). Romidepsin decreases mitochondrial membrane potential which results in reduced mitochondrial function and increased generation of reactive oxygen species and that contributes to DNA damage and apoptosis[44]. Sensitivity to the combination of romidepsin and carboplatin, but not with either single agent alone, may be explained by both high protein abundance of romidepsin's designated targets and the observation of highly enriched mitochondrial and OXPHOS pathways in chemotherapy resistant PDXs (**Figure 6A**). In addition, DNA damage driven by romidepsin may enhance the cytotoxicity of platinum DNA damaging agents such as carboplatin. Future follow-up studies will be required for confirmation.

In contrast to previous preclinical studies that use positive data from a limited number of PDX models (usually 1-2 PDX models) to drive clinical trials, our study demonstrates that benefit from targeted agents alone and in combination with chemotherapy is limited in unselected populations, similar to trials that are not biomarker driven. Indeed, two out of the four chemotherapy refractory models, BCM-7841 and HCI-028, remained insensitive to all seven targeted agents alone and in combination with carboplatin. Curiously, addition of either romidepsin or EPZ011969 to carboplatin in the HCI-015 PDX showed some evidence of antagonism such that CR to carboplatin alone converted to only a PR in combination (**Figure 7F-L**) similar to the effect of added docetaxel in that model (**Figure 1A**). Furthermore, HCI-015 PDX demonstrated clear multi-agent antagonism along with other models such as BCM-4013 and FCP699 (**Figure 7A** and **Figure 7F-L**) suggesting that negative interactions with chemotherapy agents can occur with targeted agents in addition to other chemotherapy agents we previously described (**Figure 1** and **Figure 5**). In summary, our results highlight the potential of our approach to facilitate rational design of pre-clinical drug studies using PDX models.

Discussion

Cytotoxic chemotherapy regimens for TNBC patients frequently include several agents given in combination, or in series, without knowledge of whether a given agent would be effective or ineffective for individual patients. This is impossible to address in clinical trials since individual patient responses can only be evaluated to one treatment arm as it is not possible to assign the same patient to multiple arms. We addressed this critical issue by leveraging a clinically-relevant, large cohort of 50 orthotopically transplanted PDXs where each model was assessed for response to human equivalent doses of single agent carboplatin, docetaxel, or the combination. We also leveraged multiple publicly available PDX and neoadjuvant TNBC datasets in which mice or patients were treated with similar chemotherapy regimens for validation throughout the study to suggest clinical relevance of our findings, particularly if these gene panels can be refined. Our investigation into the lack of enhanced combination response compared to the best single agent for a given TNBC PDXs has numerous important biological and clinical implications discussed below.

Molecularly guided therapy is largely lacking for TNBC, and the underlying molecular associations of treatment response to platinum derivatives, taxane compounds, and their combination is poorly understood. Molecular profiling data from multiple omic platforms in baseline PDX tumors prior to treatment associated gene sets related to chromatin regulation that were upregulated in tumors with better response to single agent carboplatin vs docetaxel (**Figure 4A**). Changes in chromatin accessibility have been linked to sensitivity and resistance to DNA cross-linking, platinum chemotherapy agents such as cisplatin and these changes have been associated with impacts on the transcriptional landscape[45]. In contrast, in PDXs for which docetaxel response was better than carboplatin, both molecular data associated higher levels of genes with roles in metabolism and structural integrity (**Figure 4A** and **Figure S4B**). More specifically, metabolism genes such as Glutathione Synthetase (GSS) and Glutathione S-Transferase Kappa 1 (GSTK1) are involved in the synthesis of a detoxification enzyme, glutathione, and conjugation of glutathione to substrates promotes elimination of toxic compounds from cells. There is evidence for glutathione promoting resistance to alkylating agents including platinum agents[46]. After cisplatin enters cells, it can be complexed with glutathione and subsequently exported out of cells[46]. This may explain higher levels of GSS and GSTK1 observed in PDXs with better response to docetaxel than carboplatin where these genes may contribute to a degree of platinum insensitivity due to carboplatin being inactivated by elevated glutathione metabolism. In addition, genes positively regulating the microtubule and actin cytoskeleton were found to be higher in PDXs that responded better to docetaxel than carboplatin, suggesting a level of microtubule-actin coordination[33] which may impact responses to taxanes since their mechanism of action involves microtubule disruption. Together, these data provide biological insights into the distinct underlying associations to different chemotherapy agents.

There is a great need to identify additional biomarkers for precision diagnostics beyond ER, PR, and HER2. In contrast to neoadjuvant pCR predictors for taxane-anthracycline based regimens using gene expression[47,48], predictors for anthracycline-free and platinum-containing regimens are understudied. Incorporation of

carboplatin with taxane-anthracycline regimens improved pCR rates compared with taxane-anthracycline alone in TNBC patients from the BrighTNess trial[6] (58% vs 31%, respectively). Furthermore, multiple TNBC trials examining anthracycline-free regimens consisting only of carboplatin + taxane have shown pCR rates above 50%[10,28,49,50], suggesting opportunities for chemotherapy optimization. Although the immune-checkpoint inhibitor pembrolizumab is now approved for neoadjuvant use in combination with carboplatin/taxane/anthracycline-containing chemotherapy, this regimen provided a 7% improvement in pCR rate compared with chemotherapy alone[13], suggesting a critical need for predictive biomarkers to identify patients that will benefit from immune-checkpoint inhibition.

The high performance and high specificity of our protein-informed, multi-omic pCR predictor for platinum + taxane (**Figure 3I**) has important implications for identifying patients that will respond to chemotherapy without the need for additional pembrolizumab and anthracycline/cyclophosphamide. Integrating proteomic data with transcriptomic data demonstrates the added value of incorporating diverse molecular profiles to enhance chemotherapy response predictions. Moreover, the ProMS and ProMS^{mo} selected biomarkers that comprise our chemotherapy response predictors to single-agent platinum, taxane, and their combination form a relatively small set of genes that is tractable for further validation and clinical translation and therefore attractive for future development into a precision diagnostic. We have trained a final composite model using all available datasets for each treatment type and these are ready to be applied to any additional dataset, such as the BEAUTY[31] (NCT02022202), TBCRC 030[32] (NCT01982448), and RESPONSE (NCT05020860) clinical trials.

Our previous WGCNA/CTD work employed pseudo-bulk RNAseq data in which the mouse stromal reads were added to the human epithelial tumor cell reads in an attempt to mimic bulk RNAseq data from human-only clinical samples. Unfortunately not all datasets were appropriate for such analysis. However, going forward, using stromal genes from pseudo-bulk data could also help improve our predictions as these genes have been demonstrated to be highly informative, especially for taxane response prediction, in our previous study that used a complementary approach (WGCNA/CTD) for chemotherapy response prediction[14]. Together with a recently described high-performance 20 gene biomarker panel to predict pCR to neoadjuvant taxane plus anthracycline-based chemotherapy regimens in TNBC[48], these predictors cover the multitude of chemotherapy regimens given to TNBC patients and provides strong rationale for incorporation into future clinical trials. Models of negative biomarker combinations will also be very useful to predict progressive disease or residual cancer burden, but will require more studies that include these measurements together with molecular profiling.

Despite finding differences in molecular associations to various chemotherapy agents described above, our investigation into associations with response to any treatment arm tested identified KRT5, a basal cytokeratin associated with basal-like breast cancer, as a potential biomarker to stratify TNBC tumors for treatment. Immunohistochemistry for basal cytokeratins 5/6, 14, and 17 have been used previously to define basal-like breast cancer and examine associations with chemotherapy responses to different treatment regimens[51–54]. Further stratifying TNBC tumors using positivity for at least one of these basal cytokeratins identified patients

with better adjuvant response to cyclophosphamide, methotrexate, and 5-fluorouracil (CMF) chemotherapy[52] and to capecitabine[51] but also identified patients with worse adjuvant response to anthracycline-based chemotherapy[53,54] compared to patients without these markers. The presence of these markers was also associated with worse response of TNBC patients treated with anthracycline-based chemotherapy in the neoadjuvant setting[55]. Here, we observed high KRT5 protein abundance by mass-spectrometry-based measurements in TNBC tumors responsive to non-anthracycline-based chemotherapy consisting of single agent carboplatin, docetaxel, or their combination. These data suggest basal-like markers may be either positive or negative markers of TNBC response to chemotherapy depending on the regimen and also provide rationale for further investigation for KRT5 as a companion diagnostic. We confirmed KRT5 findings by IHC in PDX tumors and a KRT5 Allred IHC score achieved a high AUROC of 0.83 for discriminating responsive PDXs from this study to any chemotherapy (**Figure 6F**). Additional validation of KRT5 IHC in human breast cohorts with chemotherapy response information is ongoing. We also identified a group of genes in PDXs quantified at the RNA level, aided by protein level measurements, that was predictive of neoadjuvant chemotherapy response in TNBC tumors (**Figure 3**). Future diagnostic assays and prospective clinical trials will be required to evaluate the clinical utility of these candidate biomarker or biomarker combinations.

In PDXs resistant to all treatment arms we also found elevated levels of OXPHOS and lower levels of glucose metabolism/glycolysis related pathways (**Figure 6A**). These findings of high OXPHOS but low glycolysis in our resistant PDXs is consistent with a previous study that also showed elevated OXPHOS and decreased glycolysis in a residual TNBC PDX tumor post adriamycin/cyclophosphamide chemotherapy treatment that was generated from a treatment-naive TNBC patient[56]. In addition, mitochondrial gene expression and protein synthesis were among the top pathways enriched in resistant PDXs, as was OXPHOS, which may be fueled by an increase in respiratory chain complex components (**Figure 6A**). The finding that OXPHOS may also interfere with optimal response to combination treatment (**Figure 5D**) is inline with a recent study where transcriptomic profiling of TNBC tumors prior to sequential taxane and anthracycline-based neoadjuvant therapy associated an OXPHOS signature with greater risk of recurrence and also observed that pharmacological targeting of OXPHOS suppressed tumor growth of many PDX models generated from patients with residual tumors[57]. This also aligns with other studies that implicate divergent effects of platinum vs taxane agents on mitochondrial metabolism[58] and demonstrate distinct effects of platinum plus taxane combination vs single agents on mitochondrial structure[59]. The contribution of OXPHOS may therefore be critical to identify tumors insensitive to chemotherapy and/or used in treatment decisions for combination carboplatin + docetaxel and deserves further study.

Although using PDXs in our study addresses many challenges in TNBC, it faces limitations that may impact clinical translation. Immune-deficient PDX models lack an effective immune response, a critical biological process that is well-established to become altered during cancer. The presence of specific immune cells in human tumors, such as tumor infiltrating lymphocytes, have been shown to be not only predictive of response to some chemotherapeutics, but also prognostic for TNBC patients[60,61]. Another immunity-related issue is our previous observation where immunologically “cold” TNBC tumors have higher engraftment rates as PDX models

compared to immune-enriched “hot” TNBC tumors[62] leading to the vast majority of the TNBC models analyzed in the current study to be immunologically “cold”. It is not completely understood what the impact of an intact immune system would have on our study’s findings. However, we found a substantial overlap of genes associated with drug responses in tumors from PDXs and clinical trials (**Figure 2D-F**) and also found groups of genes predictive of chemotherapy response (**Figure 3**). Similarities between PDX models and patients were also highlighted in our previous study that showed highly concordant drug responses between PDX and patient-of-origin, especially for taxanes[14]. Lastly, our study investigated tumor responses after four cycles of chemotherapy similar to evaluating patient responses to neoadjuvant chemotherapy. Evaluation of longer term measures such as tumor recurrence is an ongoing area of investigation by our team.

In conclusion, our proteogenomic characterization identifies candidate molecular mechanisms underlying response and resistance, as well as putative predictive biomarkers for stratifying TNBC tumors for single or combination chemotherapy treatments. These results also support evaluation of rationally selected targeted agents and rationally selected PDX most likely to respond to such agents (or not) that may augment chemotherapy response to a meaningful extent, and provide a valuable resource for researchers and clinicians.

Acknowledgments

We honor the breast cancer patients who donates their biopsies for PDX generation. Ms Susan Rafte provided research advocacy support for this work. This work was supported by NIH/NCI grants U54 CA224076 (to M.T.L, A.L.W., and B.E.W.), U24 CA226110 (to M.T.L.), SPORE P50 CA186784 (to C.K.O. and M.T.L.), U24 CA210954 and U24 CA271076 (both to B.Z.), R37 CA269783-01A1 (to G.V.E.), K22 Career Transition Award 1K22CA241113-01 (to G.V.E.), and Cancer Center Support Grant P30 CA125123 (to C.K.O.), by a Translational Breast Cancer Research Training Program grant T32 CA203690 (to J.T.L.), and by Core Facility grants from the Cancer Prevention and Research Institute of Texas (CPRIT) RP170691 and RP220646 (both to M.T.L.), NIH 1S10OD023469 and CPRIT RP200504 (both to D.K.), and additionally by a CPRIT Recruitment of Rising Stars Award RR160027 (to B.Z.) and CPRIT First time-faculty recruitment grant RR200009 (to G.V.E.), and an American Cancer Society Research Scholar Grant RSG-22-093-01-CCB (to G.V.E.). N.Z. was supported by CPRIT RP220468. M.A. acknowledges support from U01CA214125 and NCI-SPORE P50 CA186784-06. This work was also supported by a grant from the V Foundation and a generous gift from the Korell family for the study of triple-negative breast cancer. The BCM Mass Spectrometry Proteomics Core is supported by the Dan L Duncan Comprehensive Cancer Center NIH award P30 CA125123 and a CPRIT Core Facility Award RP17005. M.T.L, B.Z., and G.V.E. are CPRIT Scholars in Cancer Research, and B.Z. is also a McNair Scholar.

Author Contributions

Conceptualization, M.T.L., B.Z., G.V.E., M.A., and J.T.L.; data collection and processing, L.E.D., A.N.L., A.B.S., A.M., C.S., N.Z., J.C., R.R.S., D.C.K., B.W., C.H., A.J., K.L., and J.T.L.; data analysis, J.T.L., M.T.L., B.Z., R.R.S., C.H., S.V.V., T.W., C.I.M., Y.L., and V.P.; Data curation, R.R.S.; validation, L.E.D. and J.D.L.; Resources, M.T.L., B.Z., S.L., M.J.E., M.F.R., S.G.H., C.K.O., G.M.W., E.M., S.D., X.Z., F.M-B., X.C., B.E.W., and A.L.W.; writing - original draft: J.T.L., M.T.L., and B.Z.; funding acquisition: M.T.L., A.L.W., B.E.W., B.Z., and C.K.O.; Supervision, M.T.L. and B.Z.

Declaration of Interests

M.T.L is a Founder of, and an uncompensated Limited Partner in, StemMed Ltd., and an uncompensated Manager in StemMed Holdings L.L.C., its General Partner. M.T.L. is also a Founder of, and equity stake holder in, Tvardi Therapeutics Inc.. L.E.D. is a compensated employee of StemMed Ltd. Selected BCM PDX models described herein are exclusively licensed to StemMed Ltd. resulting in tangible property royalties to M.T.L. and L.E.D. Washington University in St. Louis has licensed selected PDX to Envigo which results in tangible property royalties to S.L. He also received research funding from Pfizer, Takeda Oncology, Zenopharm, independent of this project. S.L has received license fees from Envigo and also received research funding from Pfizer, Takeda Oncology, Zenopharm, outside of this project. M.J.E. is founder of Progendis, Inc and received consulting fees from Abbvie, Sermonix, Pfizer, AstraZeneca, Celgene, NanoString, Puma, Veracyte, Eli Lilly and Novartis, and is an equity stockholder and Board Director member of BioClassifier. M.J.E. is an inventor on a patent for the Breast Cancer PAM50-based assay, Prosigna, which is marketed by Veracyte. M.J.E. also receives royalties from Washington University in St. Louis when the WHIM PDX models are licensed to for-profit companies. M.F.R. received consulting fees from Pfizer, Novartis, AstraZeneca, and Gilead. C.K.O. is a stockholder of GeneTex. S.D. received consulting free from Taiho Oncology. F.M-B has advisory roles in Black Diamond, Biovica, Eisai, FogPharma, Immunomedics, Inflection Biosciences, Karyopharm Therapeutics, Loxo Oncology, Mersana Therapeutics, OnCusp Therapeutics, Puma Biotechnology Inc., Seattle Genetics, Sanofi, Silverback Therapeutics, Spectrum Pharmaceuticals, Zentalis; and received consulting fee from AbbVie, Aduro BioTech, Alkermes, AstraZeneca, Daiichi Sankyo, DebioPharm, Ecor1 Capital, eFFECTOR Therapeutics, F. Hoffman-La Roche, GT Apeiron, Genentech, Harbinger Health, IBM Watson, Infinity Pharmaceuticals, Jackson Laboratory, Kolon Life Science, Lengo Therapeutics, Menarini Group, OrigiMed, PACT Pharma, Parexel International, Pfizer, Protai Bio, Samsung Bioepis, Seattle Genetics, Tallac Therapeutics, Tyra Biosciences, Xencor, Zymeworks, European Organization for Research and Treatment of Cancer (EORTC), European Society for Medical Oncology (ESMO). B.Z. received consulting fee from AstraZeneca. G.V.E. receives sponsored research funding from Chimerix, Inc, and experimental research compounds from Chimerix, Inc and the Lead Discovery Center of Germany. M.A. received research funding from AstraZeneca. University of Utah may license the HCI PDX models described herein to for-profit companies, which may result in tangible property royalties to A.L.W. and/or B.E.W. A.L.W. has received research funding from AbbVie. All other authors have nothing to disclose.

Supplemental Figure Legends

Figure S1. Combination carboplatin and docetaxel is largely ineffective at generating enhanced responses over the best single agent in TNBC PDXs. Related to Figure 1. A) Three preclinical trials from which the 50 TNBC PDXs models in the "chemotherapy response cohort" were chosen. **B)** Spider plots depicting tumor volumes during the course of treatment ($n \geq 3$ mice per treatment arm for each PDX model). **C)** Barplots depicting computed expected combination PDX tumor responses from the Bliss independence model (see Methods) and observed combination PDX tumor responses. P-values determined by paired t-test and Wilcoxon signed rank test. No statistical difference indicates that the single agents are additive.

Figure S2. Molecular associates of carboplatin, docetaxel, and combination treatment. Related to Figure 2. A-C) Association of individual omic platforms with tumor responses to carboplatin (A), docetaxel (B), and combination (C). With WES data, plots were generated showing significant enrichment ($p < 0.05$) of deleterious variants and high-level amplification events (GISTIC = +2) associated with responsive (blue dots) or resistant (red dots) PDXs with p-values and odds ratios derived from Fisher's Exact Test (dashed horizontal line indicates $p = 0.05$). No deep deletion events (GISTIC = –2) were found to be significantly associated with any treatment arm and therefore not shown. With RNA-Seq and proteomics data, plots were generated showing significantly associated genes ($p < 0.05$) that are higher in PDXs with increasing responsiveness to treatment (blue dots) or higher in PDXs with increasing resistance to treatment (red dots) with p-values derived from Spearman's correlation between RNA/protein measurements and quantitative $\log_2$ fold change in tumor responses.

Figure S3. Predicting complete response to chemotherapy. Related to Figure 3. A) Monte-Carlo Cross Validation procedure used for training and evaluating predictive models for chemotherapy response. **B)** Area under ROC curve (AUROC) performance in the set-aside cross-validation (CV) data from the same cohort of trained logistic regression models repeated using ProMS feature selection (5 features) for predicting pathologic complete response. Numbers indicate average AUROC from 50 times of Monte-Carlo cross-validations. Datasets used are summarized in **Table S4B**. **C-F)** PCA plots depicting platinum (C), taxane (D), and platinum + taxane combination (E-F) datasets with RNA (B-D) and protein (E) data before and after merging datasets post batch correction with ComBat. **G)** Additional TNBC PDX models profiled by RNAseq and their chemotherapy responses. CR: complete response, PR: partial response, SD: stable disease, PD: progressive disease. **H)** Performance of composite platinum and taxane predictors trained on all existing data assessed using PDX models described in (G) and from RNAseq data generated from baseline TNBC tumors from a TNBC clinical trial with neoadjuvant responses to single-agent taxane.

Figure S4. Molecular features at baseline discriminate best response to single agent carboplatin or docetaxel. Related to Figure 4. A) Association of individual omic platforms with PDXs that showed better response to single agent carboplatin or docetaxel. PDX group stratification shown in **Figure 1A** for "Q2". With

WES data, plots were generated showing significant enrichment ($p < 0.05$) of deleterious variants and high-level amplification events (GISTIC = +2) associated with PDXs where response to docetaxel > carboplatin (blue dots) or PDXs where response to carboplatin > docetaxel (red dots) with p-values and odds ratios derived from Fisher's Exact Test (dashed horizontal line indicates $p = 0.05$). No deep deletion events (GISTIC = –2) were found to be significantly associated with any treatment arm and therefore not shown. With RNA-Seq and proteomics data, plots were generated showing significantly associated genes ($p < 0.05$) higher in PDXs where response to docetaxel > carboplatin (blue dots) or higher in PDXs where response to carboplatin > docetaxel (red dots) with p-values derived from Wilcoxon rank-sum tests. **B)** Plot depicting any gene found to be significantly associated with a given treatment arm in more than one omic platform from (A). **C)** Heatmap depicting AUROC of genes found to be significantly higher by RNA-Seq and 1 or more additional omic platforms in PDXs whose response was better to either single-agent carboplatin or docetaxel from (B) that were tested for discriminatory power by AUROC analysis in clinical neoadjuvant TNBC datasets.

Figure S5. Actin dynamics, trafficking, and OXPHOS metabolism may underlie positive and negative interactions between single agents. Related to Figure 5. A) Alternative two-group statistical approach to **Figure 5A** to identify associated genes by RNA-Seq and proteomics with positive vs negative interactions between agents. **B-C)** Results from approach (A) with plots showing significantly associated genes by RNA-Seq and proteomics where genes are lower in combo better (blue dots in (B)), higher in combo better (red dots in (B)), lower in combo worse (blue dots in (C)), and higher in combo worse (red dots in (C)) with with p-values derived from Wilcoxon rank-sum tests. **D)** Heatmap showing top 25 genes that are significantly higher and lower in combo better PDXs as determined by signed (indicating directionality of fold change) $-\log_{10}$ meta-p-values derived from combining p-values from genes quantified by both RNA-Seq and proteomics data in (B). Signed (indicating directionally of fold change) $-\log_{10}$ p-values and significance for individual omic platform associations. **E)** Similar to (D) except top 25 genes that are significantly higher and lower in combo worse PDXs are shown.

Figure S6. KRT5 is a chemotherapy response marker for carboplatin, docetaxel, and their combination. Related to Figure 6. A) Association of individual omic platforms with PDXs that had response to any treatment arm (PR or better) and those that were resistant to all treatment arms (mean $\log_2$ fold change or 95% CI in SD or PD range). PDX group stratification shown in **Figure 1A** for "Q4". With WES data, plots were generated showing significant enrichment ($p < 0.05$) of deleterious variants and high-level amplification events (GISTIC = +2) associated with PDXs that had response to any treatment arm (blue dots) or PDXs resistant to all treatment arms (red dots) with p-values and odds ratios derived from Fisher's Exact Test (dashed horizontal line indicates $p = 0.05$). No deep deletion events (GISTIC = –2) were found to be significantly associated with any treatment arm and therefore not shown. With RNA-Seq and proteomics data, plots were generated showing significantly associated genes ($p < 0.05$) higher in PDXs with a response to any treatment arm (blue dots) or higher in PDXs resistant to all treatment arms (red dots) with p-values derived from Wilcoxon rank-sum tests. **B)** Boxplots depicting RNA and protein levels of KRT5, KRT6A and KRT17 in baseline tumors from patients in CPTAC-TNBC trials with pathological complete response (pCR) vs non-pCR to neoadjuvant carboplatin + docetaxel

combination. P-values derived from Wilcoxon rank-sum test and Anderson-Darling test. **C)** Plot depicting AUROC using KRT5 RNA or protein levels from CPTAC-TNBC study[10] to discriminate non-pCR tumors from pCR tumors.

Supplemental Table Titles and Legends

Table S1. Breast PDX enrollment from three preclinical trials and chemotherapy response and metadata, related to Figures 1 and S1.

Table S2. Processed genomic, transcriptomic, and proteomic data associated with this study, related to Figures 2 and S2.

Table S3. Association results for mutation, copy number, RNA-Seq, and proteomics with single agent and combination treatment, related to Figures 2 and S2.

Table S4. External datasets used for chemotherapy response association and prediction and predictive biomarker combinations, related to Figures 2, 3, 4, S3, and S4.

Table S5. Gene and pathway associations for discriminating best single agent response, related to Figures 4 and S4.

Table S6. Gene and pathway associations for positive and negative interactions between single agents, related to Figures 5 and S5.

Table S7. Gene and pathway associations for resistant vs responsive PDXs, related to Figures 6 and S6.

References

1. Engel, C., Rhiem, K., Hahnen, E., Loibl, S., Weber, K.E., Seiler, S., Zachariae, S., Hauke, J., Wappenschmidt, B., Waha, A., et al. (2018). Prevalence of pathogenic BRCA1/2 germline mutations among 802 women with unilateral triple-negative breast cancer without family cancer history. BMC Cancer *18*, 265.

2. Bear, H.D., Anderson, S., Brown, A., Smith, R., Mamounas, E.P., Fisher, B., Margolese, R., Theoret, H., Soran, A., Wickerham, D.L., et al. (2003). The effect on tumor response of adding sequential preoperative docetaxel to preoperative doxorubicin and cyclophosphamide: preliminary results from National Surgical Adjuvant Breast and Bowel Project Protocol B-27. J. Clin. Oncol. *21*, 4165–4174.

3. Mamounas, E.P., Bryant, J., Lembersky, B., Fehrenbacher, L., Sedlacek, S.M., Fisher, B., Wickerham, D.L., Yothers, G., Soran, A., and Wolmark, N. (2005). Paclitaxel after doxorubicin plus cyclophosphamide as adjuvant chemotherapy for node-positive breast cancer: results from NSABP B-28. J. Clin. Oncol. *23*, 3686–3696.

4. Lehmann, B.D., Bauer, J.A., Chen, X., Sanders, M.E., Chakravarthy, A.B., Shyr, Y., and Pietenpol, J.A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J. Clin. Invest. *121*, 2750–2767.

5. von Minckwitz, G., Schneeweiss, A., Loibl, S., Salat, C., Denkert, C., Rezai, M., Blohmer, J.U., Jackisch, C., Paepke, S., Gerber, B., et al. (2014). Neoadjuvant carboplatin in patients with triple-negative and HER2-positive early breast cancer (GeparSixto; GBG 66): a randomised phase 2 trial. Lancet Oncol. *15*, 747–756.

6. Loibl, S., O'Shaughnessy, J., Untch, M., Sikov, W.M., Rugo, H.S., McKee, M.D., Huober, J., Golshan, M., von Minckwitz, G., Maag, D., et al. (2018). Addition of the PARP inhibitor veliparib plus carboplatin or carboplatin alone to standard neoadjuvant chemotherapy in triple-negative breast cancer (BrighTNess): a randomised, phase 3 trial. Lancet Oncol. *19*, 497–509.

7. Sikov, W.M., Berry, D.A., Perou, C.M., Singh, B., Cirrincione, C.T., Tolaney, S.M., Kuzma, C.S., Pluard, T.J., Somlo, G., Port, E.R., et al. (2015). Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant once-per-week paclitaxel followed by dose-dense doxorubicin and cyclophosphamide on pathologic complete response rates in stage II to III triple-negative breast cancer: CALGB 40603 (Alliance). J. Clin. Oncol. *33*, 13–21.

8. Zheng, R., Han, S., Duan, C., Chen, K., You, Z., Jia, J., Lin, S., Liang, L., Liu, A., Long, H., et al. (2015). Role of taxane and anthracycline combination regimens in the management of advanced breast cancer: a meta-analysis of randomized trials. Medicine *94*, e803.

9. Zhang, L., Wu, Z.-Y., Li, J., Lin, Y., Liu, Z., Cao, Y., Zhang, G., Gao, H.-F., Yang, M., Yang, C.-Q., et al. (2022). Neoadjuvant docetaxel plus carboplatin vs epirubicin plus cyclophosphamide followed by docetaxel in triple-negative, early-stage breast cancer (NeoCART): Results from a multicenter, randomized controlled, open-label phase II trial. Int. J. Cancer *150*, 654–662.

10. Anurag, M., Jaehnig, E.J., Krug, K., Lei, J.T., Bergstrom, E.J., Kim, B.-J., Vashist, T.D., Huynh, A.M.T., Dou, Y., Gou, X., et al. (2022). Proteogenomic Markers of Chemotherapy Resistance and Response in Triple-Negative Breast Cancer. Cancer Discov. *12*, 2586–2605.

11. Symmans, W.F., Wei, C., Gould, R., Yu, X., Zhang, Y., Liu, M., Walls, A., Bousamra, A., Ramineni, M., Sinn, B., et al. (2017). Long-Term Prognostic Risk After Neoadjuvant Chemotherapy Associated With Residual Cancer Burden and Breast Cancer Subtype. J. Clin. Oncol. *35*, 1049–1060.

12. Schmid, P., Cortes, J., Pusztai, L., McArthur, H., Kümmel, S., Bergh, J., Denkert, C., Park, Y.H., Hui, R., Harbeck, N., et al. (2020). Pembrolizumab for early triple-negative breast cancer. N. Engl. J. Med. *382*,

810–821.

13. Pusztai, L., Denkert, C., O'Shaughnessy, J., Cortes, J., Dent, R., McArthur, H., Kümmel, S., Bergh, J., Park, Y.H., Hui, R., et al. (2024). Event-free survival by residual cancer burden with pembrolizumab in early-stage TNBC: exploratory analysis from KEYNOTE-522. Ann. Oncol. *35*, 429–436.

14. Petrosyan, V., Dobrolecki, L.E., Thistlethwaite, L., Lewis, A.N., Sallas, C., Srinivasan, R.R., Lei, J.T., Kovacevic, V., Obradovic, P., Ellis, M.J., et al. (2023). Identifying biomarkers of differential chemotherapy response in TNBC patient-derived xenografts with a CTD/WGCNA approach. iScience *26*, 105799.

15. Kent, D.M., and Hayward, R.A. (2007). Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. JAMA *298*, 1209–1212.

16. Zhang, X., Claerhout, S., Prat, A., Dobrolecki, L.E., Petrovic, I., Lai, Q., Landis, M.D., Wiechmann, L., Schiff, R., Giuliano, M., et al. (2013). A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. Cancer Res. *73*, 4885–4897.

17. Meric-Bernstam, F., Lloyd, M.W., Koc, S., Evrard, Y.A., McShane, L.M., Lewis, M.T., Evans, K.W., Li, D., Rubinstein, L., Welm, A., et al. (2024). Assessment of patient-derived xenograft growth and antitumor activity: The NCI PDXNet consensus recommendations. Mol. Cancer Ther. *23*, 924–938.

18. Bliss, C.I. (1956). The calculation of microbial assays. Bacteriol. Rev. *20*, 243–258.

19. Huang, D., Savage, S.R., Calinawan, A.P., Lin, C., Zhang, B., Wang, P., Starr, T.K., Birrer, M.J., and Paulovich, A.G. (2021). A highly annotated database of genes associated with platinum resistance in cancer. Oncogene *40*, 6395–6405.

20. Savage, S.R., Yi, X., Lei, J.T., Wen, B., Zhao, H., Liao, Y., Jaehnig, E.J., Somes, L.K., Shafer, P.W., Lee, T.D., et al. (2024). Pan-cancer proteogenomics expands the landscape of therapeutic targets. Cell. https://doi.org/10.1016/j.cell.2024.05.039.

21. Sheils, T.K., Mathias, S.L., Kelleher, K.J., Siramshetty, V.B., Nguyen, D.-T., Bologa, C.G., Jensen, L.J., Vidović, D., Koleti, A., Schürer, S.C., et al. (2021). TCRD and Pharos 2021: mining the human proteome for disease biology. Nucleic Acids Res. *49*, D1334–D1346.

22. Savage, P., Pacis, A., Kuasne, H., Liu, L., Lai, D., Wan, A., Dankner, M., Martinez, C., Muñoz-Ramos, V., Pilon, V., et al. (2020). Chemogenomic profiling of breast cancer patient-derived xenografts reveals targetable vulnerabilities for difficult-to-treat tumors. Commun Biol *3*, 310.

23. Silver, D.P., Richardson, A.L., Eklund, A.C., Wang, Z.C., Szallasi, Z., Li, Q., Juul, N., Leong, C.-O., Calogrias, D., Buraimoh, A., et al. (2010). Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer. J. Clin. Oncol. *28*, 1145–1153.

24. 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 anti-cancer potential of non-oncology drugs by systematic viability profiling. Nat Cancer *1*, 235–248.

25. Wolf, D.M., Yau, C., Wulfkuhle, J., Brown-Swigart, L., Gallagher, R.I., Lee, P.R.E., Zhu, Z., Magbanua, M.J., Sayaman, R., O'Grady, N., et al. (2022). Redefining breast cancer subtypes to guide treatment prioritization and maximize response: Predictive biomarkers across 10 cancer therapies. Cancer Cell *40*, 609–623.e6.

26. Li, W., Newitt, D.C., Gibbs, J., Wilmes, L.J., Jones, E.F., Arasu, V.A., Strand, F., Onishi, N., Nguyen, A.A.-T., Kornak, J., et al. (2020). Predicting breast cancer response to neoadjuvant treatment using multi-feature MRI: results from the I-SPY 2 TRIAL. NPJ Breast Cancer *6*, 63.

27. Gonzalez-Angulo, A.M., Akcakanat, A., Liu, S., Green, M.C., Murray, J.L., Chen, H., Palla, S.L., Koenig, K.B., Brewster, A.M., Valero, V., et al. (2014). Open-label randomized clinical trial of standard neoadjuvant

chemotherapy with paclitaxel followed by FEC versus the combination of paclitaxel and everolimus followed by FEC in women with triple receptor-negative breast cancer. Ann. Oncol. *25*, 1122–1127.

28. Menghi, F., Banda, K., Kumar, P., Straub, R., Dobrolecki, L., Rodriguez, I.V., Yost, S.E., Chandok, H., Radke, M.R., Somlo, G., et al. (2022). Genomic and epigenomic BRCA alterations predict adaptive resistance and response to platinum-based therapy in patients with triple-negative breast and ovarian carcinomas. Sci. Transl. Med. *14*, eabn1926.

29. Shi, Z., Wen, B., Gao, Q., and Zhang, B. (2021). Feature Selection Methods for Protein Biomarker Discovery from Proteomics or Multiomics Data. Mol. Cell. Proteomics *20*, 100083.

30. Chang, J.C., Wooten, E.C., Tsimelzon, A., Hilsenbeck, S.G., Gutierrez, M.C., Elledge, R., Mohsin, S., Osborne, C.K., Chamness, G.C., Allred, D.C., et al. (2003). Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet *362*, 362–369.

31. Goetz, M.P., Kalari, K.R., Suman, V.J., Moyer, A.M., Yu, J., Visscher, D.W., Dockter, T.J., Vedell, P.T., Sinnwell, J.P., Tang, X., et al. (2017). Tumor sequencing and patient-derived xenografts in the neoadjuvant treatment of breast cancer. J. Natl. Cancer Inst. *109*. https://doi.org/10.1093/jnci/djw306.

32. Mayer, E.L., Abramson, V., Jankowitz, R., Falkson, C., Marcom, P.K., Traina, T., Carey, L., Rimawi, M., Specht, J., Miller, K., et al. (2020). TBCRC 030: a phase II study of preoperative cisplatin versus paclitaxel in triple-negative breast cancer: evaluating the homologous recombination deficiency (HRD) biomarker. Ann. Oncol. *31*, 1518–1525.

33. Goode, B.L., Drubin, D.G., and Barnes, G. (2000). Functional cooperation between the microtubule and actin cytoskeletons. Curr. Opin. Cell Biol. *12*, 63–71.

34. Vanhoefer, U., Harstrick, A., Wilke, H., Schleucher, N., Walles, H., Schröder, J., and Seeber, S. (1995). Schedule-dependent antagonism of paclitaxel and cisplatin in human gastric and ovarian carcinoma cell lines in vitro. Eur. J. Cancer *31A*, 92–97.

35. Xia, L., Xiao, X., Liu, W.L., Song, Y., Liu, T.J.J., Li, Y.J., Zacksenhaus, E., Hao, X.J., and Ben-David, Y. (2018). Coactosin-like protein CLP/Cotl1 suppresses breast cancer growth through activation of IL-24/PERP and inhibition of non-canonical TGFβ signaling. Oncogene *37*, 323–331.

36. Hirokawa, N., Niwa, S., and Tanaka, Y. (2010). Molecular motors in neurons: transport mechanisms and roles in brain function, development, and disease. Neuron *68*, 610–638.

37. Yadav, N., Kumar, S., Marlowe, T., Chaudhary, A.K., Kumar, R., Wang, J., O'Malley, J., Boland, P.M., Jayanthi, S., Kumar, T.K.S., et al. (2015). Oxidative phosphorylation-dependent regulation of cancer cell apoptosis in response to anticancer agents. Cell Death Dis. *6*, e1969.

38. Schott, A.F., Landis, M.D., Dontu, G., Griffith, K.A., Layman, R.M., Krop, I., Paskett, L.A., Wong, H., Dobrolecki, L.E., Lewis, M.T., et al. (2013). Preclinical and clinical studies of gamma secretase inhibitors with docetaxel on human breast tumors. Clin. Cancer Res. *19*, 1512–1524.

39. Krug, K., Jaehnig, E.J., Satpathy, S., Blumenberg, L., Karpova, A., Anurag, M., Miles, G., Mertins, P., Geffen, Y., Tang, L.C., et al. (2020). Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. Cell *183*, 1436–1456.e31.

40. Gou, X., Kim, B.-J., Anurag, M., Lei, J.T., Young, M.N., Holt, M.V., Fandino, D., Vollert, C.T., Singh, P., Alzubi, M.A., et al. (2023). Kinome reprogramming is a targetable vulnerability in ESR1 fusion-driven breast cancer. Cancer Res. *83*, 3237–3251.

41. Zhao, N., Cao, J., Xu, L., Tang, Q., Dobrolecki, L.E., Lv, X., Talukdar, M., Lu, Y., Wang, X., Hu, D.Z., et al. (2018). Pharmacological targeting of MYC-regulated IRE1/XBP1 pathway suppresses MYC-driven breast cancer. J. Clin. Invest. *128*, 1283–1299.

42. Nair, A., Chung, H.-C., Sun, T., Tyagi, S., Dobrolecki, L.E., Dominguez-Vidana, R., Kurley, S.J., Orellana, M., Renwick, A., Henke, D.M., et al. (2018). Combinatorial inhibition of PTPN12-regulated receptors leads to a broadly effective therapeutic strategy in triple-negative breast cancer. Nat. Med. *24*, 505–511.

43. Gao, Y., Kabotyanski, E.B., Shepherd, J.H., Villegas, E., Acosta, D., Hamor, C., Sun, T., Montmeyor-Garcia, C., He, X., Dobrolecki, L.E., et al. (2021). Tumor suppressor PLK2 may serve as a biomarker in triple-negative breast cancer for improved response to PLK1 therapeutics. Cancer Res. Commun. *1*, 178–193.

44. Valdez, B.C., Brammer, J.E., Li, Y., Murray, D., Liu, Y., Hosing, C., Nieto, Y., Champlin, R.E., and Andersson, B.S. (2015). Romidepsin targets multiple survival signaling pathways in malignant T cells. Blood Cancer J. *5*, e357.

45. Wang, L., Peng, Q., Yin, N., Xie, Y., Xu, J., Chen, A., Yi, J., Tang, J., and Xiang, J. (2021). Chromatin accessibility regulates chemotherapy-induced dormancy and reactivation. Mol. Ther. Nucleic Acids *26*, 269–279.

46. Ishikawa, T., and Ali-Osman, F. (1993). Glutathione-associated cis-diamminedichloroplatinum(II) metabolism and ATP-dependent efflux from leukemia cells. Molecular characterization of glutathione-platinum complex and its biological significance. J. Biol. Chem. *268*, 20116–20125.

47. Edlund, K., Madjar, K., Lebrecht, A., Aktas, B., Pilch, H., Hoffmann, G., Hofmann, M., Kolberg, H.-C., Boehm, D., Battista, M., et al. (2021). Gene expression-based prediction of neoadjuvant chemotherapy response in early breast cancer: Results of the prospective multicenter EXPRESSION trial. Clin. Cancer Res. *27*, 2148–2158.

48. Inayatullah, M., Mahesh, A., Turnbull, A.K., Dixon, J.M., Natrajan, R., and Tiwari, V.K. (2024). Basal-epithelial subpopulations underlie and predict chemotherapy resistance in triple-negative breast cancer. EMBO Mol. Med. *16*, 823–853.

49. Sharma, P., López-Tarruella, S., García-Saenz, J.A., Khan, Q.J., Gómez, H.L., Prat, A., Moreno, F., Jerez-Gilarranz, Y., Barnadas, A., Picornell, A.C., et al. (2018). Pathological response and survival in triple-negative breast cancer following neoadjuvant carboplatin plus docetaxel. Clin. Cancer Res. *24*, 5820–5829.

50. Sharma, P., López-Tarruella, S., García-Saenz, J.A., Ward, C., Connor, C.S., Gómez, H.L., Prat, A., Moreno, F., Jerez-Gilarranz, Y., Barnadas, A., et al. (2017). Efficacy of neoadjuvant carboplatin plus docetaxel in triple-negative breast cancer: Combined analysis of two cohorts. Clin. Cancer Res. *23*, 649–657.

51. Zhu, Y., Li, K., Zhang, J., Wang, L., Sheng, L., and Yan, L. (2021). The prognostic and predictive significance of cytokeratin 5/6 and epidermal growth factor receptor in metastatic triple-negative breast cancer treated with maintenance capecitabine. Transl. Cancer Res. *10*, 1193–1203.

52. Rakha, E.A., Elsheikh, S.E., Aleskandarany, M.A., Habashi, H.O., Green, A.R., Powe, D.G., El-Sayed, M.E., Benhasouna, A., Brunet, J.-S., Akslen, L.A., et al. (2009). Triple-negative breast cancer: distinguishing between basal and nonbasal subtypes. Clin. Cancer Res. *15*, 2302–2310.

53. Conforti, R., Boulet, T., Tomasic, G., Taranchon, E., Arriagada, R., Spielmann, M., Ducourtieux, M., Soria, J.C., Tursz, T., Delaloge, S., et al. (2007). Breast cancer molecular subclassification and estrogen receptor expression to predict efficacy of adjuvant anthracyclines-based chemotherapy: a biomarker study from two randomized trials. Ann. Oncol. *18*, 1477–1483.

54. Cheang, M.C.U., Voduc, D., Bajdik, C., Leung, S., McKinney, S., Chia, S.K., Perou, C.M., and Nielsen, T.O. (2008). Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. Clin. Cancer Res. *14*, 1368–1376.

55. Abdelrahman, A.E., Rashed, H.E., Abdelgawad, M., and Abdelhamid, M.I. (2017). Prognostic impact of

EGFR and cytokeratin 5/6 immunohistochemical expression in triple-negative breast cancer. Ann. Diagn. Pathol. *28*, 43–53.

56. Echeverria, G.V., Ge, Z., Seth, S., Zhang, X., Jeter-Jones, S., Zhou, X., Cai, S., Tu, Y., McCoy, A., Peoples, M., et al. (2019). Resistance to neoadjuvant chemotherapy in triple-negative breast cancer mediated by a reversible drug-tolerant state. Sci. Transl. Med. *11*. https://doi.org/10.1126/scitranslmed.aav0936.

57. Evans, K.W., Yuca, E., Scott, S.S., Zhao, M., Paez Arango, N., Cruz Pico, C.X., Saridogan, T., Shariati, M., Class, C.A., Bristow, C.A., et al. (2021). Oxidative Phosphorylation Is a Metabolic Vulnerability in Chemotherapy-Resistant Triple-Negative Breast Cancer. Cancer Res. *81*, 5572–5581.

58. Baek, M.L., Lee, J., Pendleton, K.E., Berner, M.J., Goff, E.B., Tan, L., Martinez, S.A., Mahmud, I., Wang, T., Meyer, M.D., et al. (2023). Mitochondrial structure and function adaptation in residual triple negative breast cancer cells surviving chemotherapy treatment. Oncogene *42*, 1117–1131.

59. Berner, M.J., Beasley, H.K., Vue, Z., Lane, A., Vang, L., Baek, M.L., Marshall, A.G., Killion, M., Zeleke, F., Shao, B., et al. (2024). Three-dimensional analysis of mitochondria in a patient-derived xenograft model of triple negative breast cancer reveals mitochondrial network remodeling following chemotherapy treatments. bioRxivorg. https://doi.org/10.1101/2024.09.09.611245.

60. Loi, S., Sirtaine, N., Piette, F., Salgado, R., Viale, G., Van Eenoo, F., Rouas, G., Francis, P., Crown, J.P.A., Hitre, E., et al. (2013). Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. J. Clin. Oncol. *31*, 860–867.

61. Wang, K., Xu, J., Zhang, T., and Xue, D. (2016). Tumor-infiltrating lymphocytes in breast cancer predict the response to chemotherapy and survival outcome: A meta-analysis. Oncotarget *7*, 44288–44298.

62. Petrosyan, V., Dobrolecki, L.E., LaPlante, E.L., Srinivasan, R.R., Bailey, M.H., Welm, A.L., Welm, B.E., Lewis, M.T., and Milosavljevic, A. (2022). Immunologically “cold” triple negative breast cancers engraft at a higher rate in patient derived xenografts. NPJ Breast Cancer *8*, 104.

63. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. *50*, D543–D552.

64. Harvey, J.M., Clark, G.M., Osborne, C.K., and Allred, D.C. (1999). Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J. Clin. Oncol. *17*, 1474–1481.

65. Saltzman, A.B., Leng, M., Bhatt, B., Singh, P., Chan, D.W., Dobrolecki, L., Chandrasekaran, H., Choi, J.M., Jain, A., Jung, S.Y., et al. (2018). gpGrouper: A Peptide Grouping Algorithm for Gene-Centric Inference and Quantitation of Bottom-Up Proteomics Data. Mol. Cell. Proteomics *17*, 2270–2283.

66. Evrard, Y.A., Srivastava, A., Randjelovic, J., Doroshow, J.H., Dean, D.A., 2nd, Morris, J.S., Chuang, J.H., and NCI PDXNet Consortium (2020). Systematic Establishment of Robustness and Standards in Patient-Derived Xenograft Experiments and Analysis. Cancer Res. *80*, 2286–2297.

67. Conway, T., Wazny, J., Bromage, A., Tymms, M., Sooraj, D., Williams, E.D., and Beresford-Smith, B. (2012). Xenome--a tool for classifying reads from xenograft samples. Bioinformatics *28*, i172–i178.

68. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv [q-bio.GN]. https://doi.org/10.48550/ARXIV.1303.3997.

69. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-

generation DNA sequencing data. Nat. Genet. *43*, 491–498.

70. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

71. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly *6*, 80–92.

72. Cingolani, P., Patel, V.M., Coon, M., Nguyen, T., Land, S.J., Ruden, D.M., and Lu, X. (2012). Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front. Genet. *3*, 35.

73. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443.

74. Landrum, M.J., Lee, J.M., Benson, M., Brown, G.R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., et al. (2018). ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. *46*, D1062–D1067.

75. Kuilman, T., Velds, A., Kemper, K., Ranzani, M., Bombardelli, L., Hoogstraat, M., Nevedomskaya, E., Xu, G., de Ruiter, J., Lolkema, M.P., et al. (2015). CopywriteR: DNA copy number detection from off-target sequence data. Genome Biol. *16*, 49.

76. Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhim, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. *12*, R41.

77. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

78. Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc. *11*, 2301–2319.

79. Jeffrey T. Leek and W. Evan Johnson and Hilary S. Parker and Elana J. Fertig and Andrew E. Jaffe and Yuqing Zhang and John D. Storey and Leonardo Collado Torres (2024). sva: Surrogate Variable Analysis. https://doi.org/10.18129/B9.bioc.sva.

80. Liao, Y., Wang, J., Jaehnig, E.J., Shi, Z., and Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res. *47*, W199–W205.

81. Savage, S.R., Shi, Z., Liao, Y., and Zhang, B. (2019). Graph Algorithms for Condensing and Consolidating Gene Set Analysis Results. Mol. Cell. Proteomics *18*, S141–S152.

82. Dowd, C. (2022). twosamples: Fast Permutation Based Two Sample Tests. Preprint.

83. Corsello, S.M., Bittker, J.A., Liu, Z., Gould, J., McCarren, P., Hirschman, J.E., Johnston, S.E., Vrcic, A., Wong, B., Khan, M., et al. (2017). The Drug Repurposing Hub: a next-generation drug library and information resource. Nat. Med. *23*, 405–408.

Methods

RESOURCE AVAILABILITY

Lead Contact

Further information and requests should be directed to and will be fulfilled by the lead author, Michael. T. Lewis (mtlewis@bcm.edu)

Materials Availability

Xenografts are available from the lead author (mtlewis@bcm.edu) for academic/nonprofit use on a cost recovery basis via a Material Transfer Agreement.

Data and code availability

Genomic data have been deposited in the sequence read archive identifier PRJNA756268. RNA-Seq data have been deposited with the NCBI gene expression omnibus (GEO) accession number GSE183187. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[63] partner repository with the dataset identifier PXD035857. Sample annotations, processed and normalized data files are provided in **Table S1B** and **Table S2**. Software and code used in this study are referenced in their corresponding STAR Methods and also the key resources table. Publicly available datasets are as described in **Table S4**. All other data are available upon reasonable request to Lead Contact.

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Patient recruitment

Patients with breast cancer were recruited from clinics in the Lester and Sue Smith Breast Center at Baylor College of Medicine (BCM) (Houston, TX, USA) and Ben Taub General Hospital (Houston, TX, USA) under Institutional Review Board-approved protocols. Most patients received initial core-needle biopsies at the time of diagnosis and again either during or after treatment.

Generation and chemotherapy response of PDX models

All human tissues were processed in compliance with NIH regulations and institutional guidelines, reviewed and approved by the Institutional Animal Care and Use Committee at Baylor College of Medicine. Single agent carboplatin, docetaxel, and untreated/vehicle control data were derived from three separate preclinical trials (**Figure S1A**) with further details as described previously in Petrosyan et al[14]. Response to chemotherapy was assessed by tumor volumes measured at pre- and post-treatment timepoints. Log_2 transformed fold change in tumor volume at 28 days after initiation of treatment was calculated. If the tumor completely disappeared after

treatment, a tumor volume of 0.05 was used in the calculation to avoid log transformation error. After $\log_2$ transformation, a general linear model within each PDX was computed to compare the differences between treatment groups, and to estimate the average $\log_2$ fold change in tumor size and its associated 95% confidence interval for each treatment within a PDX. These data were also plotted for visual presentation. Response was classified based on modified RECIST 1.1 criteria[16,17]: Complete response (CR) was defined as non-palpable tumors. Partial response (PR) was defined as PDXs with ≥ 30% reduction in tumor volume but not reaching the non-palpable state. Stable disease (SD) was defined in PDXs with < 30% decrease but no more than 20% over baseline tumor volume. Progressive disease (PD) was defined for PDXs with tumor volume increases of > 20% over baseline tumor volume. Thresholds for quantitative responses were also determined. For example, since partial response (PR) is defined as ≥ 30% decrease in tumor volume below baseline at day 28 of treatment, a PDX model with an average $\log_2$ fold change in tumor size of < -0.515 [30% shrinkage ~ $\log_2(0.7)$] would be deemed PR.

Drug combination effect using Bliss independence model

The Bliss independence model[18] was used to evaluate drug combination response. This model assumes the combined effect of two drugs in the product of their individual effects. The expected response of carboplatin + docetaxel combination for each PDX model was computed as $A + B - (A \times B)$ where A is the $\log_2$ fold change response to carboplatin alone and B is the $\log_2$ fold change response to docetaxel alone. A lower limit of −12 was given to models whose expected response was lower than −12 since this value was the lowest possible change based on palpable tumor size (see above). A paired t-test and paired Wilcoxon signed rank test was then performed to test for significant differences between computed expected vs observed $\log_2$ combination tumor response values. No statistical difference was found and therefore the combination response is interpreted as additive between the two single agents.

Immunohistochemistry

PDX tumors were harvested, formalin-fixed for 24 hours, and then preserved in 70% ethanol. Tissue was paraffin embedded, cut into 3 μm sections, and immunohistochemical staining was performed. During IHC, antigen retrieval was performed using 0.1M citrate buffer pH 6.0 and sections were incubated with mouse anti-cytokeratin 5 monoclonal antibody (MA5-12596, Invitrogen) at 1:20 dilution for 1 hour at room temperature. Sections were incubated with HRP anti-mouse antibody (DAKO) and detected with DAB+ solution (DakoCytomation) and DAB

sparkle enhancer (Biocare). Counterstain was performed using Harris modified hematoxylin. Images were acquired using a Nikon Eclipse Ci microscope with DS-Fi3 camera and NIS-Elements software. Staining for cytokeratin 5 was scored using the Allred method as previously described[64].

METHOD DETAILS

Whole exome sequencing (WES)

DNA sequencing was performed by core facilities at BCM and Cornell University. Exome Sequencing Libraries were prepared using the Agilent SureSelect XT v6.0 Human kit. Sequencing was performed on an Illumina HiSeq 4000 machine (PE, 2x100 cycles, ~133X estimated coverage per sample) at Cornell. At BCM, sequencing was performed on the Illumina NovaSeq machine (PE, 2x100, ~200M read pairs per sample). Separation of human and mouse reads, alignment to the reference genome, variant calling, and annotation were performed using the PDXNet Tumor-Only Variant Calling pipeline developed by the Jackson Laboratory and hosted on the Cancer Genomics Cloud.

RNA sequencing

Total RNA was extracted from PDX samples and 10 ng was used to generate and amplify transcriptome cDNA (NuGen Ovation v2, NuGen Technologies). Three grams of cNDA from each sample was fragmented to 250-400 bases using the Covaris S2 focused ultrasonicator (Covaris). Using the Illumina TruSeq or NovaSeq DNA-Seq library preparation kits (Illumina Technologies), a double-stranded DNA library was generated with 1 g of the sheared cDNA. The library was quantified with a Kapa quantitative PCR Library Quantification Kit (Kapa). DNA libraries (11 pM) were loaded onto an Illumina HiSeq 2000 or NovaSeq6000 FlowCell and clusters were generated on the Illumina cBot. Paired-end 100-bp reads were generated on a HiSeq 2000 or NovaSeq6000 Sequencing System.

Mass spectrometry methods

The tissues were pulverized in homemade kapton tube bags under liquid nitrogen to form powder and directly digested in 50 mM ammonium bicarbonate solution using trypsin enzyme at 37^0C overnight. The digest was acidified with 10% formic acid (FA) and peptides were measured using the Pierce™ Quantitative Colorimetric Peptide Assay (Thermo Scientific 23275). 50 µg peptides were subjected to manual offline fractionation using a high pH reversed-phase chromatography to form a 5 fraction pool of peptides as described earlier[65]. LC-MS/MS analysis was carried out using a nano-LC 1200 system (Thermo Fisher Scientific, San Jose, CA) coupled to Orbitrap Fusion™ Lumos ETD mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The fractionated peptides (1 µg) were loaded on a 2 cm X 100 µm I.D. trap-column (Reprosil-Pur Basic C18 1.9 µm beads, Dr.Maisch GmbH, Germany), eluted using a 75 min gradient of different concentration of buffer B (90% acetonitrile in 0.1% FA) as follows: [3% to 28%B - 70 min, 95%B - 5 min] at a flow rate of 750 nL/min on a 5 cm x 150 µm I.D. analytical column (Reprosil-Pur Basic C18 1.9 µm, Dr.Maisch GmbH, Germany). The peptides

were directly electro-sprayed at 2.2 kV voltage into the mass spectrometer operated in a data-dependent mode with 'top 30' method. The full MS scan was acquired in Orbitrap in the range of 300-1400 m/z at 120,000 resolution (AGC 5E5, maxIT 50 ms) followed by MS2 in Ion Trap (HCD 28% collision energy, AGC 5E3, maxIT 35 ms) with 15 sec dynamic exclusion time.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using R unless explained otherwise.

Genomic data analysis

Deleterious variant calling

All raw FASTQ files were subjected to QC verification by FASTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and were trimmed for adapter sequences with TrimGalore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Whole Exome FASTQ files were processed using the Tumor-Only Variant Calling pipeline[66] developed by the Jackson Laboratory for PDXNet hosted on the Cancer Genomics Cloud. This pipeline uses Xenome[67] to separate human epithelial reads and mouse stromal reads, then aligns human reads against the GRCh38 human genome using BWA[68]. It then uses GATK MuTect2[69,70] to call variants and SnpEff/SnpSift[71,72] to annotate variants, generating per-sample VCF and tabular outputs for downstream analyses. Further annotation with gnomAD[73] and CLINVAR[74] was performed which added information by matching locus fields (Chromosome, Position, Reference and Alternate alleles) to standard annotation VCF files. Multi-allelic sites were decomposed to ensure Alternate alleles matched exactly with annotation sources.

Copy number calling

A set of 6 normal breast tissue samples were subjected to Whole Exome Sequencing both at Cornell and Baylor to serve as platform-matched normal samples. These platform-matched normal samples were run through the Variant Calling pipeline to obtain platform-matched normal BAMs. CopywriteR[75] was used on each tumor BAM file obtained from the Variant Calling pipeline, and a random platform-matched normal BAM was used as control. CopywriteR uses depth information from off-target reads to calculate segment-level copy number data. This segment level data was then processed using GISTIC2.0[76] to obtain raw and threshold gene-level and focal-level copy number information.

RNA-Seq quantification

RNA-Seq FASTQ files were processed using Xenome to separate murine stromal reads and human epithelial reads, which were both then quantified with rsem-calculate-expression. A reference index was created for RSEM (v1.3.0)[77] using rsem-prepare-reference on the hg38 and mm10 genome assembly (FASTA) and transcript feature (GTF) files to serve as library input to the rsem-calculate-expression tool. Expected counts from RSEM were collated to create a genes-by-samples matrix. This matrix was Upper-Quantile normalized per-sample to

account for differences in sequencing depth and then log_2 transformed to reduce expression variance across orders of magnitude.

Proteomic data processing

Raw MS/MS data were processed using MaxQuant (version 1.6.5.0)[78] with the Andromeda search engine3 against the human plus mouse RefSeq database. The standard MaxQuant contaminant database was also included in the searching. Enzyme specificity was set to trypsin and up to two missed cleavages were allowed. Oxidation (M), acetyl (protein N-term) and deamidation (NQ) were set as variable modifications. No fixed modification was selected. The maximum number of modifications per peptide was set as 4. The match between runs function was enabled. A false discovery rate cutoff of 1% was used at the PSM and protein levels. Reverse and contaminant matches were removed from the analysis. The file "evidence.txt" generated by MaxQuant was processed using gpGrouper[65] to generate gene level quantification data. Different batch and normalization methods were evaluated using OmicsEV (https://github.com/bzhanglab/OmicsEV) which selected the following procedure to generate proteomic data for all downstream analyses: Gene products which were identified and quantified in at least two replicates of any PDX model were retained. Expression values were averaged across all replicates of each PDX model for which a measurement was made then batch correction was applied using ComBat from the sva R package[79] across the cohort stratified by year of sample processing.

Proteogenomic analysis

Publicly available datasets for molecular profiles of TNBC PDX and TNBC tumors

Molecular profiling data of baseline TNBC tumors prior to neoadjuvant chemotherapy was retrieved as follows. Microarray data for GSE18864[23] and GSE194040[25] was downloaded from GEO, log_2 transformed, and collapsed to gene level by mean probe expression. RNA-Seq data for TNBC PDXs from GSE142767[22] was also retrieved from GEO. RNA-Seq and Tandem Mass Tag (TMT) mass spectrometry-based proteomics data for CPTAC-TNBC trials was downloaded from supplemental tables in the Anurag et al, 2022 manuscript[10].

Molecular features and drug response associations

For associations with deleterious variants and copy number events, PDXs were binarized into responsive (CR + PR) and resistant (SD + PD) models. Fisher's exact test was performed to test for enrichment of deleterious variants in responsive or resistant models. Fisher's exact test was also performed to test for enrichment of amplification (GISTIC = +2) or deep deletion (GISTIC = −2) events in responsive or resistant models. For both mutation and copy number data, Inf and –Inf values were set to 5 and –5, respectively. Spearman's correlation was computed for genes quantified in at least 50% of PDXs by RNA-Seq and proteomics data with quantitative tumor measurements after treatment. For pathway analysis, signed $-log_{10}$ p-values (positive sign if positively

correlated with response, negative sign if negative correlated with response) from correlation analyses were used as input for gene set enrichment analysis (GSEA) with gene sets in Gene Ontology: Biological Processes, Reactome, and Wikipathway databases using WebGestalt[80]. A weighted set cover was then applied to reduce the number of genesets from each database to a minimal set while maximizing gene coverage[81]. Wilcoxon-rank sum and Anderson-Darling tests (using ad_test from the twosamples R package[82]) were performed for two group comparisons where indicated.

mRNA association with drug response in human TNBC datasets

Human TNBC tumors from public datasets with mRNA profiling were classified as responsive or resistant to treatment as follows. For GSE18864, CR + PR tumors were considered responsive and SD + PD tumors were considered resistant. For BrighTNess and CPTAC-TNBC trials, pCR tumors were considered responsive and all else were considered resistant. Wilcoxon-rank sum test was performed to assess differences between responsive vs resistant tumors in all these datasets. Effect sizes were computed from the difference in median values between the two groups. P-values less than 0.01 or 0.05 were considered significant as described in the text.

Chemotherapy response prediction with ProMS

Protein Marker Selection (ProMS) tool[29] was implemented in the python package *proms* (https://pypi.org/project/proms) to select 5 non-redundant and therefore complementary RNA markers to predict CR/pCR using mRNA alone or in combination with and protein data. TNBC samples from PDXs and clinical trial samples with mRNA or both mRNA and protein data were categorized to be responsive (CR or pCR) or resistant (all else) (**Table S4B**). Logistic regression models were first trained using individual datasets to select k = 5 markers based on molecular data using mRNA data alone and evaluated by Monte-Carlo Cross Validation (MCCV) process 50 times. Each time, the dataset is randomly split 70% for classifier training (T) and 30% for validation (V). During classifier training, a number of hyperparameters were tuned, including univariate filtering and L2 regularization through a grid search values of 0.001, 0.01, 0.1, 1, 10, and 100 with 3-fold cross-validation within the training set, T. The trained classifiers were then evaluated with the validation set, V (different for each of the 50 repeats). The same procedure was performed using ComBat[79] batch corrected mRNA merged datasets by treatment type (**Figures S3B-D**) to select k = 5-10 markers and evaluated by the MCCV process 25 times. The number of markers selected giving the highest average MCCV AUROC was selected to repeat the feature selection and classifier building process using all data to produce a final fully trained, fixed model whose performance was assessed by AUROC on independent test data to predict pCR on clinical datasets. For the platinum + taxane predictor, proteins quantified in all samples were used to generate ComBat[79] batch corrected merged protein data (**Figure S3F**) and used together with merged mRNA data (**Figure S3E**) to select 10 protein-informed, mRNA markers for logistic regression model training as described above. All other parameters were used as default for model training and testing, the details of which can be found in the ProMS manuscript[29].

Chemotherapy prediction with WGCNA/CTD

Weighted co-expression network analysis followed by connect the dots (WGCNA/CTD) was implemented in R as documented in GitHub (https://github.com/vpetrosyan/biomarkers_chemotherapy) on merged mRNA datasets (**Figures S3B-D**). Extreme Gradient Boosting (XGBoost) models using features selected by WGCNA/CTD were generated and evaluated by leave-one-out cross-validation (LOOCV) to predict pCR in clinical datasets that were independent of datasets used for WGCNA/CTD feature selection as documented in **Figures S3F-G**.

Predicting drug target vulnerabilities to targeted agents

RNA-Seq data from PDXs and primary breast tumors in the CPTAC-BRCA study[39] were used. To make tumor data comparable to PDX data, tumor mRNA profiles were scaled individually using the upper quartile of the PDX dataset. For each query PDX, its mRNA profile was queried against the CPTAC-BRCA dataset by computing Spearman's correlation coefficients between the query PDX profile and all tumor mRNA profiles in the dataset. The top 20% of CPTAC-BRCA breast tumors with highest similarities to the query PDX were considered as query-like tumors. Proteomics data was then used since proteins are the direct targets of the drugs being examined. The proteomic profiles of the query-like group were compared to those from the remaining tumors using *t*-test to identify significantly up-regulated druggable proteins as putative targets. Druggable proteins were downloaded from The Drug Repurposing Hub[83] website on September 19, 2019 then filtered by mechanism of action terms "inhibitor", "antagonist", or "blocker" to identify inhibited targets of drugs. These targets were then ranked according to p-values with the top ranked (e.g. rank 1) druggable protein being the most significantly up-regulated in the query-like group vs others.

Meta p-value calculation

Meta p-values were computed using the "sumz" method from the metap R package. P-values from assessing the difference between two groups using a Wilcoxon rank-sum test for RNA and protein were first converted to one-sided p-values. These p-values were signed based on the effect size based on the difference in medians between the two groups for each data type. If this difference was positive, then the one-sided p-value was assigned a positive sign; if the difference was negative, the one-sided p-value was assigned a negative sign. The calculated meta p-value was converted back to two-sided p-values and then the sign was taken of the majority of p-values, or in case of a tie, the sign of the most significant p-value.

Figures

Figure 1: Lack of enhanced combo response vs best single agent in TNBC PDXs

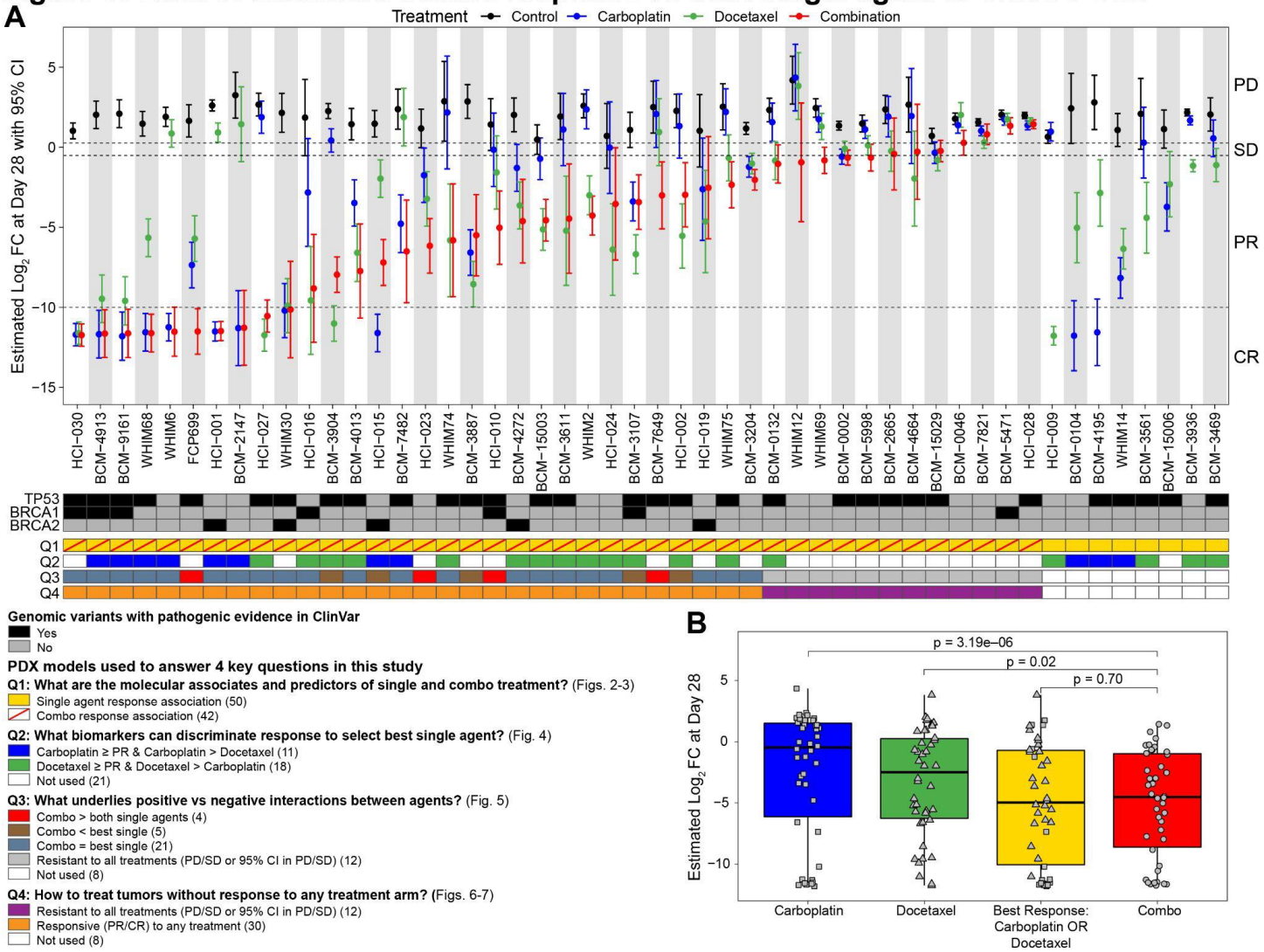


Figure 1: Combination carboplatin and docetaxel is largely ineffective at generating enhanced responses over the best single agent in TNBC PDXs. A) (Top panel) Quantitative tumor responses to chemotherapy in 50 TNBC PDX models comprising the "chemotherapy response cohort". A general linear model for each PDX was generated to estimate mean $\log_2$ fold change (FC) in tumor volume at Day 28 vs Day 0 (baseline) and its associated 95% confidence interval (CI) for each treatment within a PDX ($n \geq 3$ per treatment arm). Response was also qualitatively assessed based on modified RECIST 1.1 classification (PD, Progressive Disease; SD, Stable Disease; PR, Partial Response; CR, Complete Response). (Middle panel) Annotation of PDXs harboring pathogenic variants of TP53, BRCA1, BRCA2, or PALB2. (Bottom panel) Annotations depicting sample stratification for analyses to address four key questions (Q1 to Q4) asked in this study. **B)** For PDXs with response to all treatment arms, boxplots depict quantitative chemotherapy responses for each PDX to single agent carboplatin (blue) and docetaxel (green), the best response to either single agent carboplatin OR docetaxel (yellow), and to the combination (red). These results suggest that combination treatment does not enhance tumor shrinkage over the best single-agent in TNBC PDXs. Boxes depict the interquartile range (IQR) of the scores with horizontal lines depicting the median. Whiskers extend to 1.5 x IQR from Q1 (25th percentile) and Q3 (75th percentile), respectively. P-values derived from paired Wilcoxon signed-rank tests.

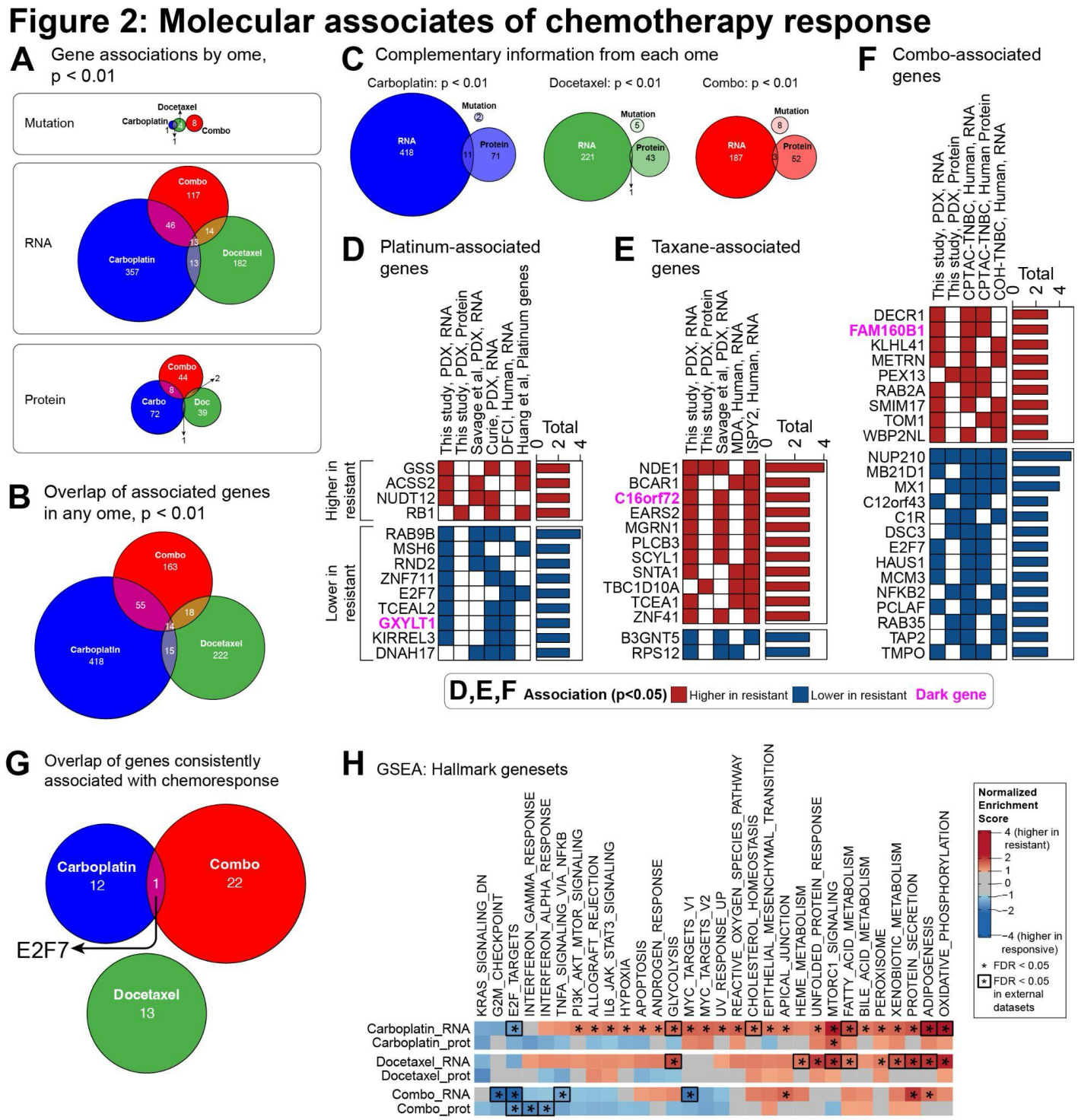


Figure 2. Molecular associates of carboplatin, docetaxel, and combination treatment. A) Venn diagrams depicting overlap of treatment-associated genes from **Figure S2**. ($p < 0.01$) by data type. **B)** Venn diagram depicting overlap of treatment-associated genes ($p < 0.01$) by any data type from **Figure S2**. **C)** Venn diagrams depicting overlap of significant ($p < 0.01$) molecular associations by treatment type. **D-F)** Plots summarizing treatment associated consensus genesets derived from this study along with external datasets for (D) platinum, (E) taxane, or (F) combination. Genes that are significant in this study's PDX dataset and also in at least two additional datasets are shown. Genes are also annotated according to a drug target tier system we developed to inform actionable insights[20] and also annotated as a dark, understudied gene from Shiels et al[21]. **G)** Venn diagram depicting overlap of genes from consensus genesets (D-F). **H)** Heatmap showing normalized enrichment scores (NES) from GSEA analysis with Hallmark genesets using RNA-Seq and proteomics data. Asterisks indicate enrichment FDR < 0.05. Outlined cells indicate significant pathways (FDR < 0.05) also found in external datasets listed in **Table S4A**.

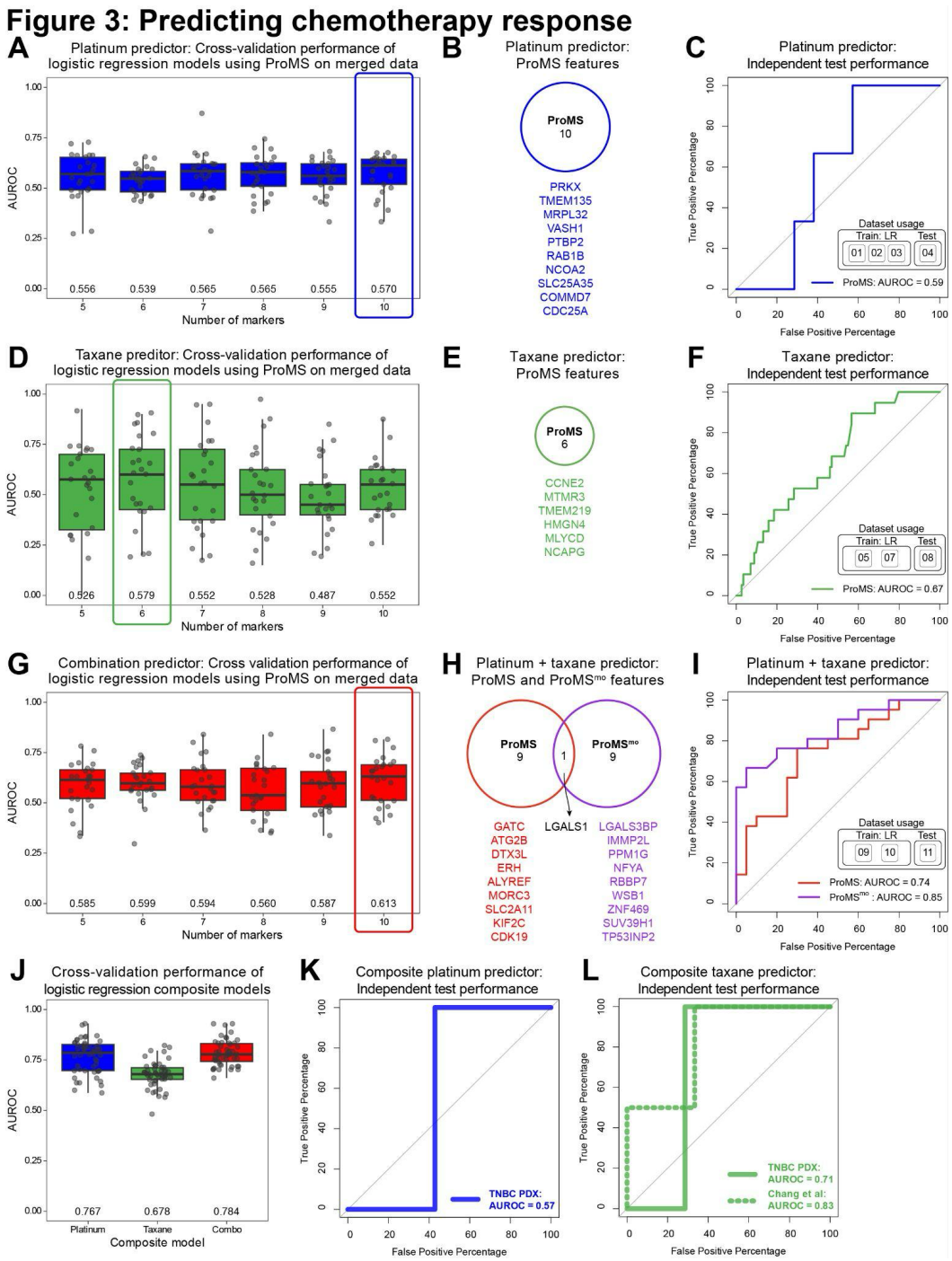


Figure 3. Predicting chemotherapy response. A,D,G) Area under the receiver operating curve (AUROC) performance in the set-aside cross-validation data repeated 25X of trained logistic regression models using ProMS feature selection (5-10 features) for predicting complete response/pathologic complete response from combined RNA datasets for single-agent platinum (A), single-agent taxane (D), and platinum + taxane combination (G). Numbers indicate average AUROC. Datasets used are summarized in **Table S4B**. **B,E,H)** ProMS features selected features giving the highest average cross-validation AUROC from (A), (D), and (G). A multi-omics, protein-facilitated RNA marker selection method (ProMS^{mo}) was also examined in (H) with 10 features which was based on the number of markers with the highest cross-validation AUROC in (G). **C,F,I)** Performance of fully trained logistic regression (LR) models using markers selected by ProMS/ProMS^{mo} in (B,E,H), respectively to predict chemoresponse in independent test data. **J)** Performance in set-aside cross-validation data repeated 50X of logistic regression models trained using all possible data (platinum datasets: 01, 02, 03, 04; taxane datasets: 05, 07, 08; combination datasets: 09, 10, 11). Numbers indicate average AUROC. **K-L)** Performance of composite platinum and taxane predictors trained on all existing data assessed using independent test datasets from PDX models described in **Figure S3G-H** to predict response to single agent platinum (K) or taxane (L).

Figure 4: Discriminating best single agent response

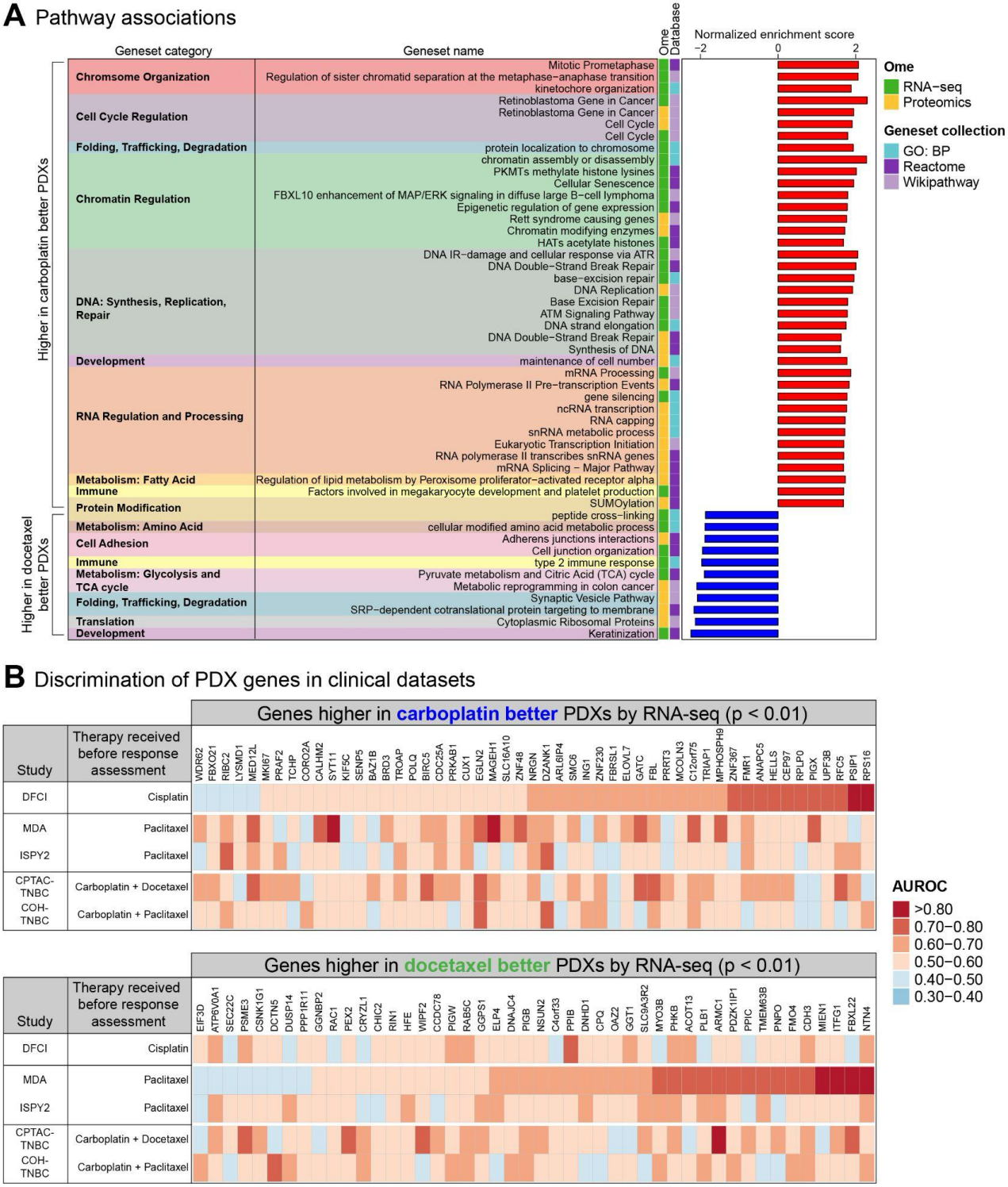


Figure 4. Molecular features at baseline discriminate best response to single agent carboplatin or docetaxel. A) GSEA analysis with Gene Ontology: Biological Processes (GO: BP), Reactome, and Wikipathway geneset collections using signed (indicating positive or negative sign of correlation coefficient) $-\log_{10}$ p-values from RNA-Seq and proteomics data in **Figure S4A**. Normalized enrichment scores of significant genesets (FDR < 0.05) are shown after applying weighted set cover to reduce geneset redundancy. Genesets were further categorized into categories by manual curation guided by leading edge genes in each geneset. **B)** Heatmaps depicting AUROC of genes found to be significantly higher by RNA-Seq data in PDXs whose response was better to either single-agent carboplatin or docetaxel from **Figure S4A** that were tested for discriminatory power by AUROC analysis in clinical neoadjuvant TNBC datasets.

Figure 5: Positive vs negative interactions between agents

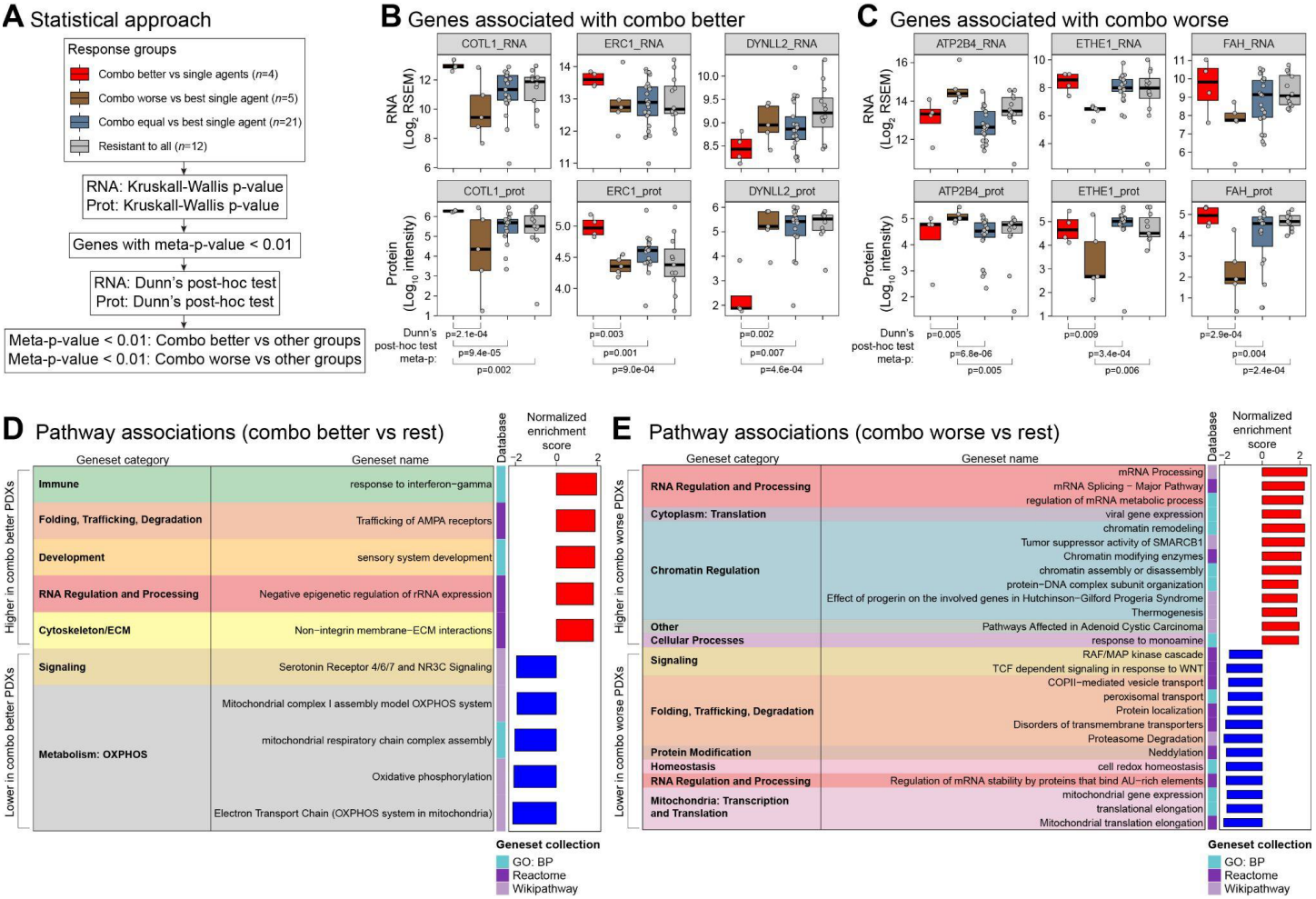


Figure 5. Actin dynamics, trafficking, and OXPHOS metabolism may underlie positive and negative interactions between single agents. A) Workflow describing multi-group comparison using Kruskal-Wallis test followed by Dunn's post-hoc test to identify differences between groups performed separately with RNA-Seq and proteomics data. Meta-p-values were generated by Stouffer's method (sumz) using post-hoc test-derived p-values from RNA-Seq and proteomics to find multi-omic support for genes associated with PDXs where combination was beneficial vs single agents or associated with PDXs where combination was worse vs best single agent. **B-C)** Boxplots showing all significant genes from (A). Genes associated with combination benefit vs single agents included genes with roles in regulating actin dynamics and trafficking (B) while genes with roles in homeostasis and metabolism associated with PDXs where combination was worse vs best single agent. Boxes depict the interquartile range (IQR) of the scores with horizontal lines depicting the median. Whiskers extend to 1.5 x IQR from Q1 (25th percentile) and Q3 (75th percentile), respectively. Meta-P-values derived as described in (A). **D-E)**. Using a complementary sample stratification and statistical approach to that described in **Figure S5A**, GSEA enrichment results (FDR < 0.05) are shown from GO: BP, Reactome, and Wikipathway gene collections using signed (indicating directionality of fold change) $-\log_{10}$ meta-p-values derived from RNA-Seq and proteomics data together. Association with cytoskeleton and homeostasis pathways support gene level results found in (A-C).

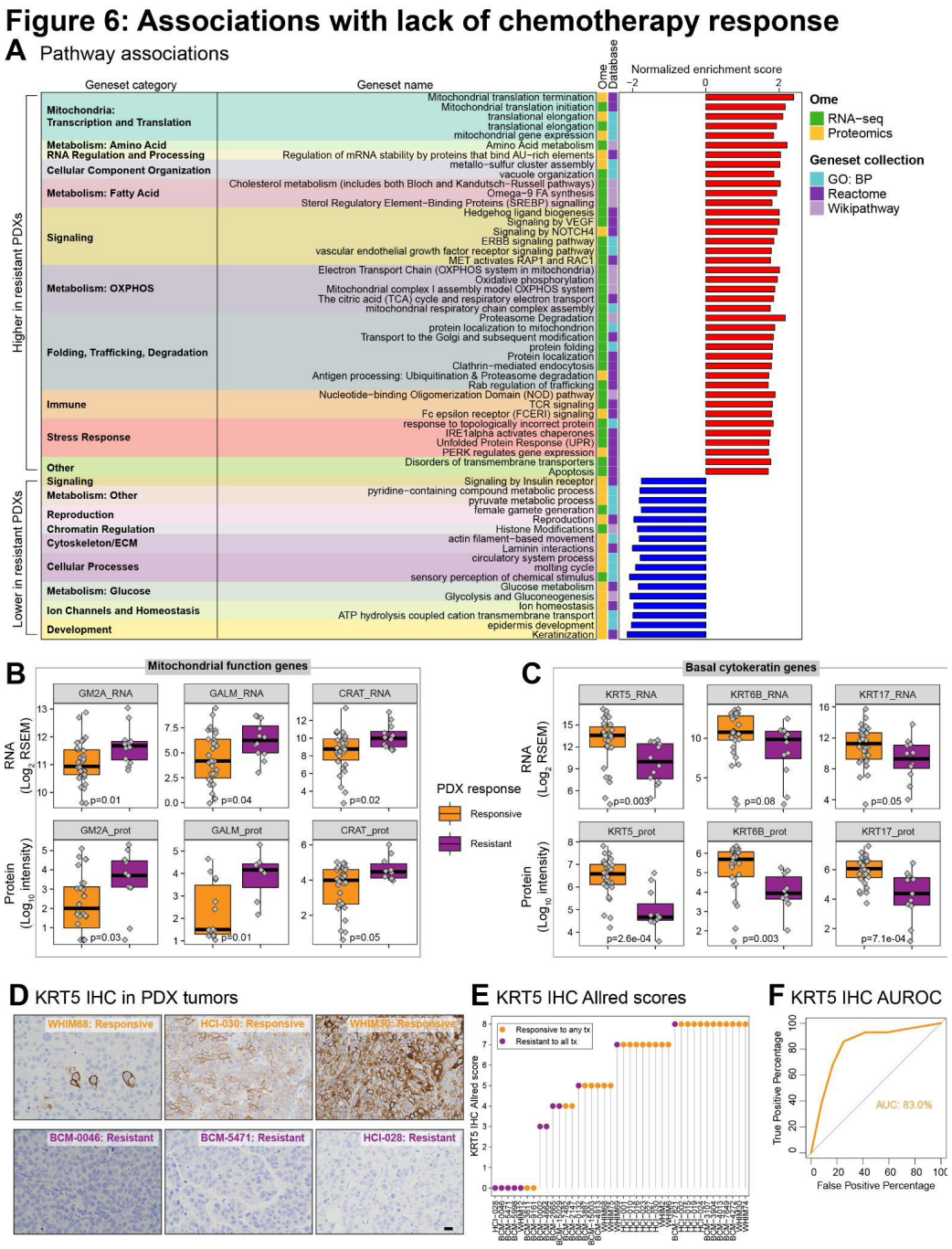


Figure 6. KRT5 is a chemotherapy response marker for carboplatin, docetaxel, and their combination. A) GSEA analysis with Gene Ontology: Biological Processes (GO: BP), Reactome, and Wikipathway geneset collections using signed (indicating positive or negative sign of correlation coefficient) $-\log_{10}$ p-values from RNA-Seq and proteomics data in **Figure S6A**. Normalized enrichment scores of significant genesets (FDR < 0.05) are shown after applying weighted set cover to reduce geneset redundancy. Genesets were further categorized into categories by manual curation guided by leading edge genes in each geneset. **B)** Boxplots depicting selected genes from **Figure S6B** that were supported by both RNA-Seq and proteomics data with roles in metabolism that are higher in PDXs resistant to all treatment arms (carboplatin, docetaxel, and the combination) compared to PDXs responsive to any treatment. **C)** Similar to (B) except showing basal marker genes that are higher in PDXs responsive to any treatment vs PDXs resistant to all treatments. For (A-B), boxes depict the interquartile range (IQR) of the scores with horizontal lines depicting the median. Whiskers extend to 1.5 x IQR from Q1 (25th percentile) and Q3 (75th percentile), respectively. P-values derived from Wilcoxon rank-sum tests. **D)** Representative images of KRT5 IHC in baseline PDX tumors prior to treatment. Scale bar = 10 μm. **E)** Plot showing Allred IHC scores for KRT5 quantified from IHC images. **F)** AUROC using KRT5 Allred scores in predicting response to any chemotherapy treatment arm for TNBC PDXs in this study.

Figure 7: Targeted agents can enhance chemotherapy response

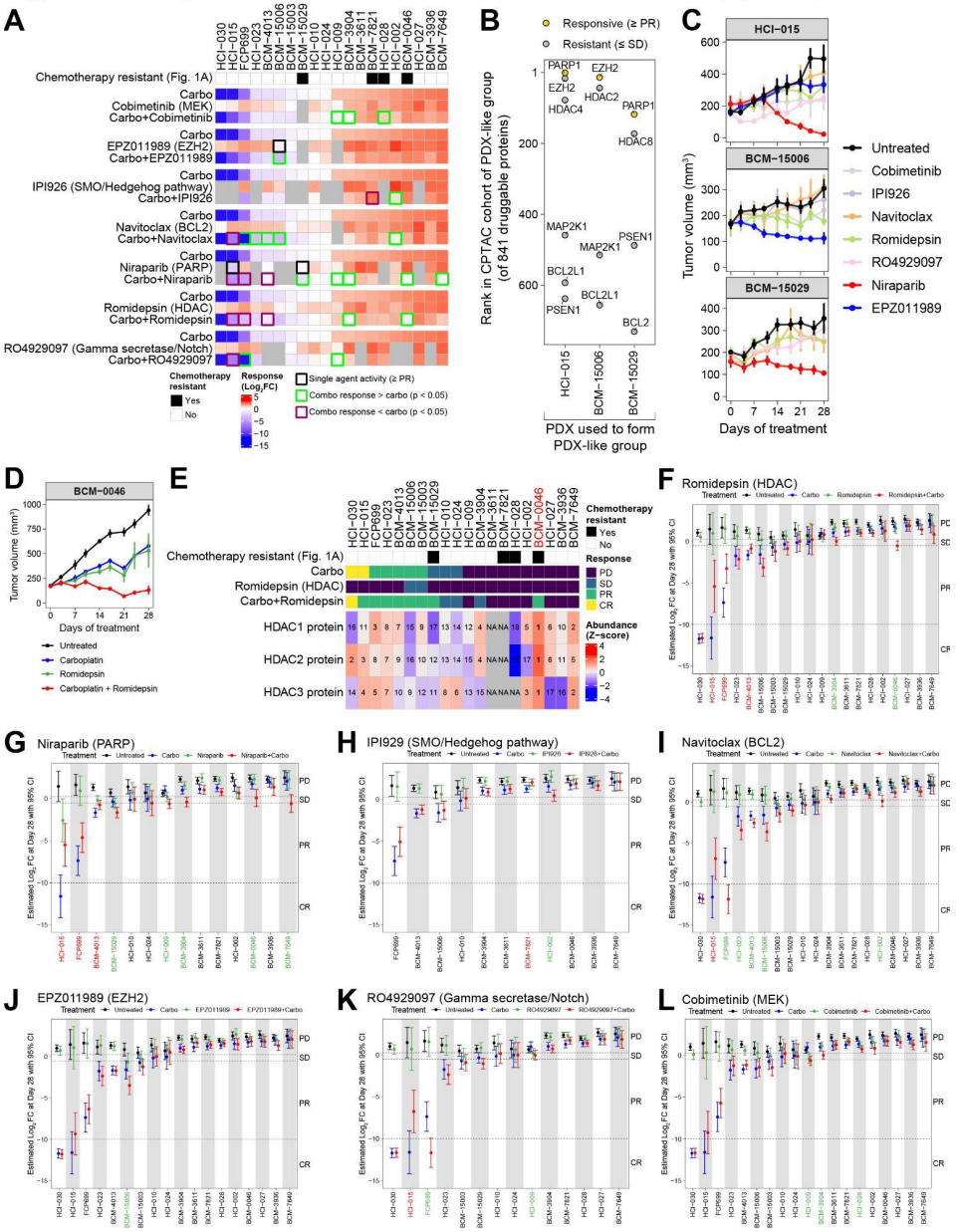


Figure 7. Targeted agents can enhance chemotherapy response. A) Quantitative responses ($\log_2$ fold change in tumor volume at day 28 vs day 0) for PDXs enrolled in the SPORE/U54 PDXNet preclinical trial (**Figure S1A**) where PDXs were treated with 7 targeted agents alone and in combination with carboplatin. Green boxes indicate PDX-drug pairs where response to the targeted agent + carboplatin combination was statistically better than carboplatin alone while red boxes indicate PDX-drug pairs where response to the targeted agent + carboplatin combination showed antagonism relative to carboplatin alone. **B)** For each of the three responsive PDXs to single agent treatment, the top 20% of CPTAC breast tumors most similar to each PDX were considered query-like tumors, and their proteomic profiles were compared to those from the remaining tumors using a *t*-test to identify significantly up-regulated druggable proteins as putative targets. These targets were then ranked from highest (most significantly upregulated) to lowest and the targets from agents in (A) were plotted. **C)** Plots showing tumor volumes of untreated controls and indicated targeted single agent treatment arms. Data shown are averages from 3-6 mice per treatment arm ± SEM. **D)** Plot showing tumor volumes of BCM-0046 for control and indicated treatment arms. Data shown are averages from 3-6 mice per treatment group ± SEM. **E)** Heatmap depicting protein abundance of romidepsin's designated targets HDAC1, HDAC2, and HDAC3 in PDX tumors. Numerals in heatmap indicate protein rank (lower value means higher abundance) for the indicated protein across all quantified PDXs. **F-L**) Plots depicting quantitative response from (A) to carboplatin alone, the indicated targeted agent alone, and in combination with carboplatin. Green and red PDX annotations indicate response to the targeted agent + carboplatin combination was statistically better or worse than carboplatin alone, respectively.

Figure S1: Lack of enhanced combo response vs best single agent

A

V Foundation Preclinical Trial

PDX "Patient" (20 PDX/arm)

N=9 Vehicle control; N=3 per treatment arm

Single Agents	Combination
Vehicle	
Docetaxel (20mg/kg)	Docetaxel + TTI-101
Carboplatin (50mg/kg)	Carboplatin + TTI-101
TTI-101 (STAT3)	

SPORE/U54 PDXNet Preclinical Trial

PDX "Patient" (14-20 PDX/arm)

N=6 Untreated control; N=3 per treatment arm

Single Agents	Combination
Untreated	
Docetaxel (20mg/kg)	
Carboplatin (50mg/kg)	Carbo + Docetaxel
EPZ011989 (EZH2)	Carbo + EPZ011989
IPI926 (SMO)	Carbo + IPI926
Navitoclax (BCL2/XL/W)	Carbo + Navitoclax
Niraparib (PARP1/2)	Carbo + Niraparib
Romidepsin (HDAC)	Carbo + Romidepsin
Cobimetinib (MEK1)	Carbo + Cobimetinib
RO4929097 (GSI/Notch)	Carbo + RO4929097

Korell Preclinical Trial

PDX "Patient" (33 PDX/arm)

N=9 Untreated control; N=3 per treatment arm

Single Agents	Combination
Untreated	
Docetaxel (20mg/kg)	Carbo + Docetaxel20
Docetaxel (30mg/kg)	Carbo + Docetaxel30
Carboplatin (50mg/kg	

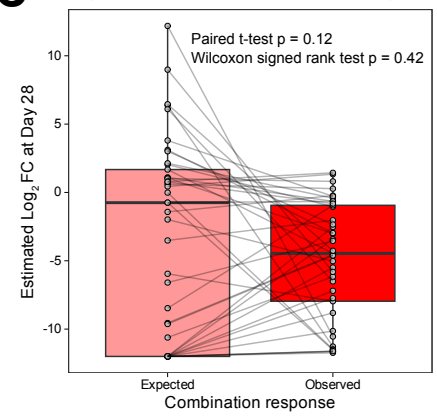
Endpoints

Tumor growth

Time to event:
1. Tumor doubling
2. Tumor halving
3. Complete response
4. Recurrence

Drug interaction

C Carboplatin + docetaxel combination response



B

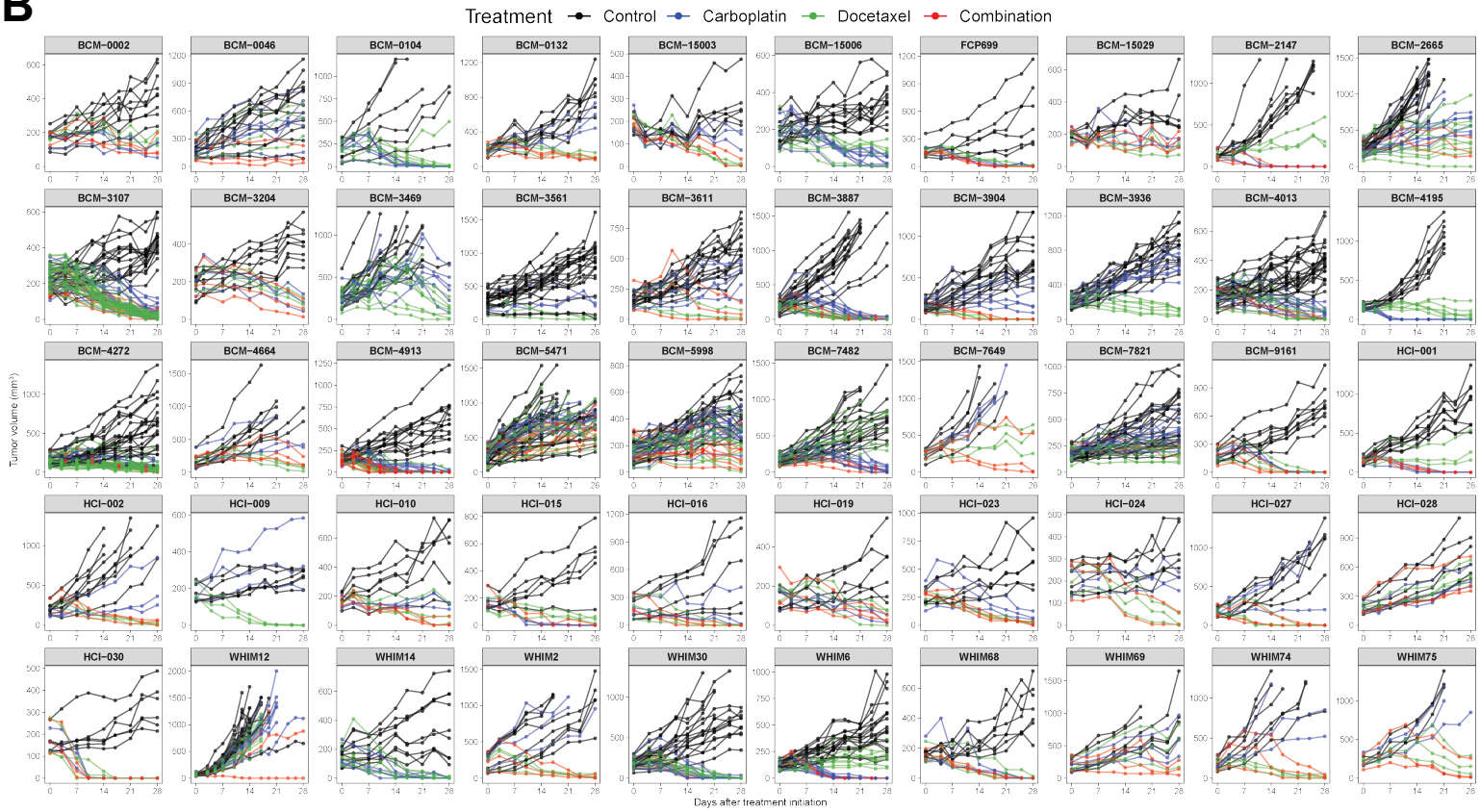


Figure S2: Molecular associates of single-agent and combination treatment

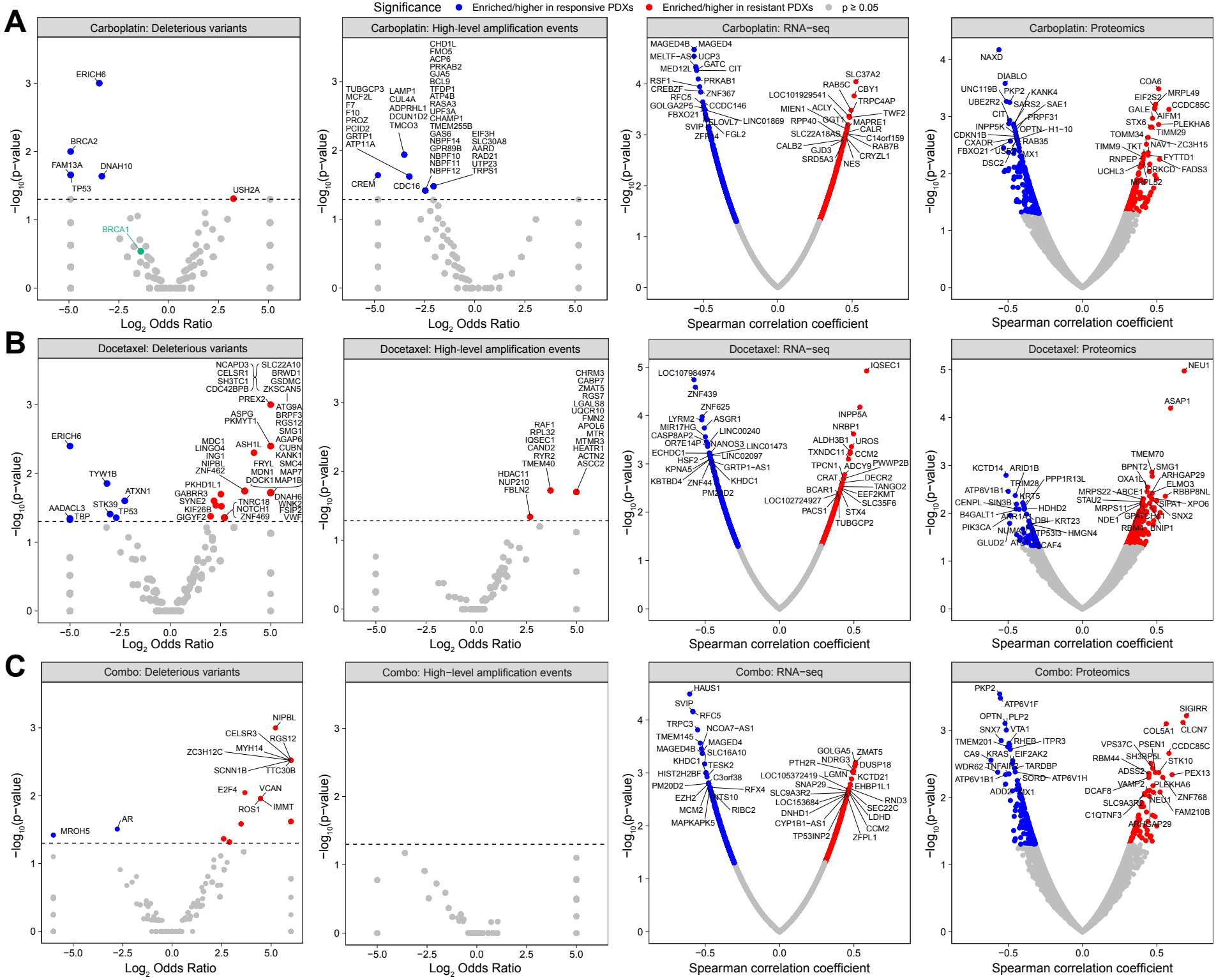
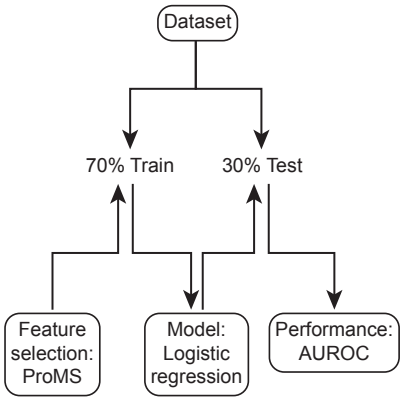
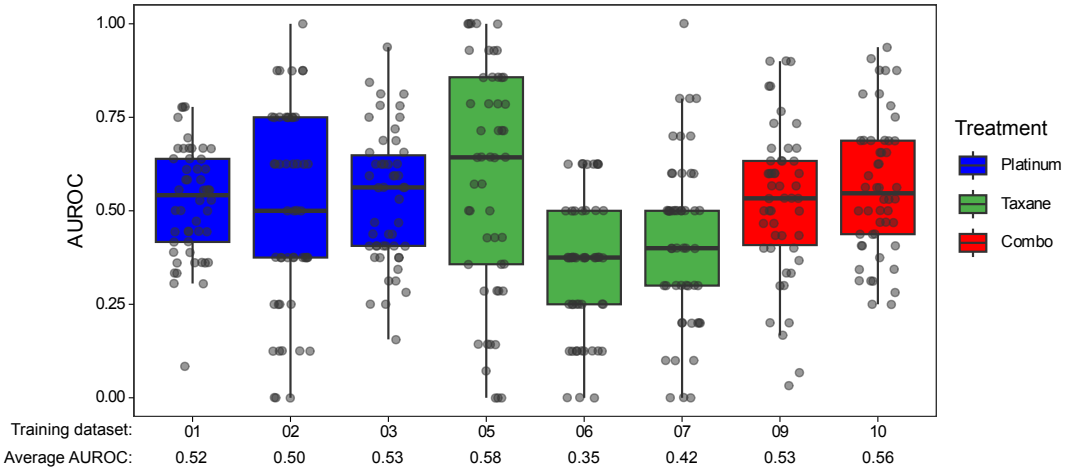


Figure S3: Predicting chemotherapy response

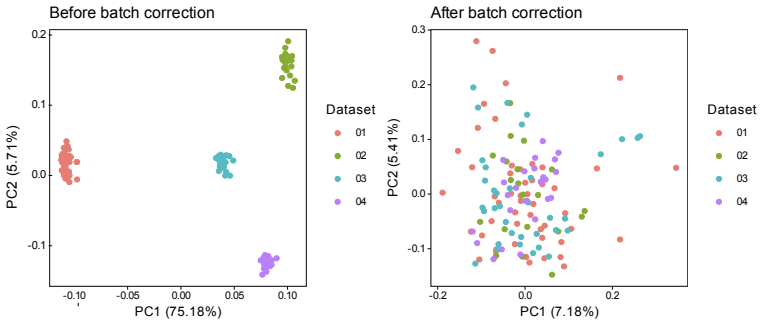
A Monte-Carlo Cross Validation



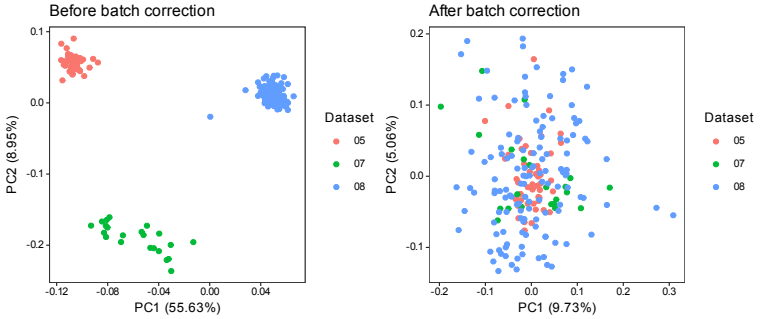
B Cross-validation performance of trained logistic regression models using 5 ProMS selected features on individual datasets



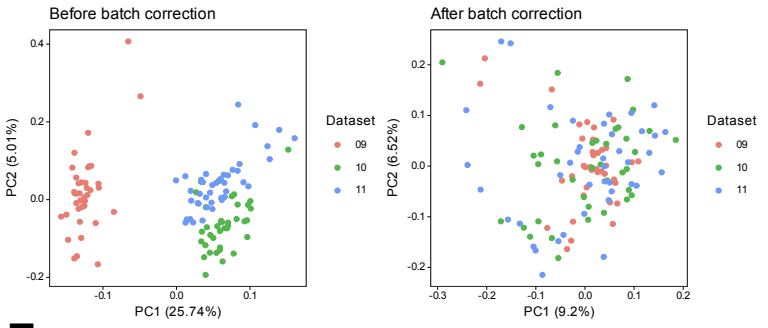
C Platinum datasets (RNA)



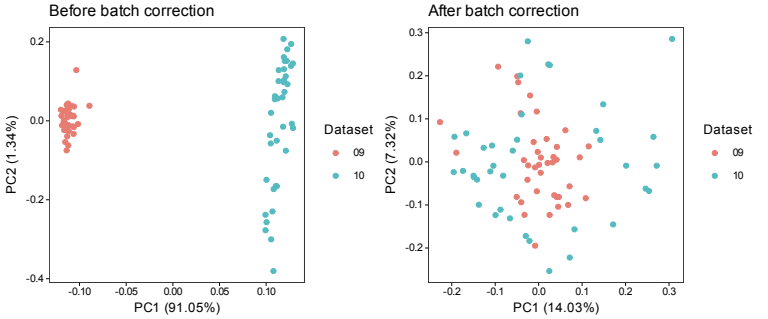
D Taxane datasets (RNA)



E Combination datasets (RNA)



F Combination datasets (Protein)



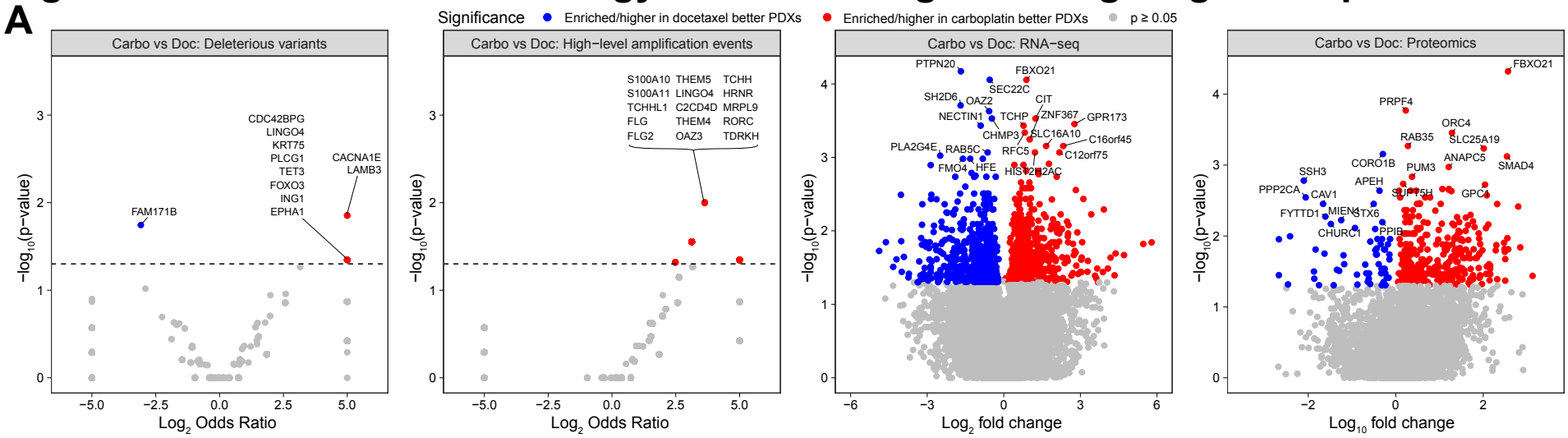
G Additional TNBC PDX models with RNAseq data

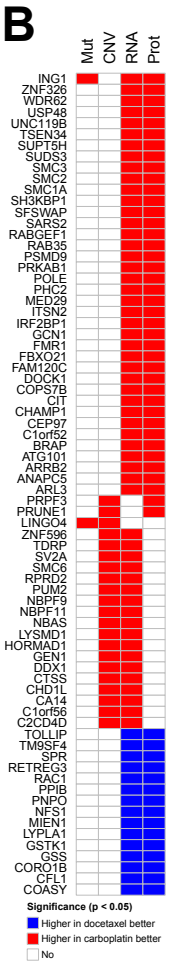
Treatment	HCI-034	BCM-15166	BCM-15116	BRA45	BCM-5438	BCM-4849	BCM-15120	BCM-2277
Control	PD	PD	PD	PD	PD	PD	PD	PD
Carboplatin	SD	PD	PR	PD	CR	SD	PR	PD
Docetaxel	PR	PR	SD	SD	PR	PD	CR	PR
Combo	PR	PR	PR	SD	PR	PR	PR	PR

H Performance of composite model with ProMS-selected features

Composite Model	TNBC RNAseq Test Dataset	Responsive samples (CR/pCR)	Resistant samples (all else)	AUROC
Platinum	TNBC PDX (Fig. S3G)	1	7	0.57
Taxane	TNBC PDX (Fig. S3G)	1	7	0.71
	Chang et al, 2003	2	9	0.83

Figure S4: Biomarkers and biology discriminating best single agent response





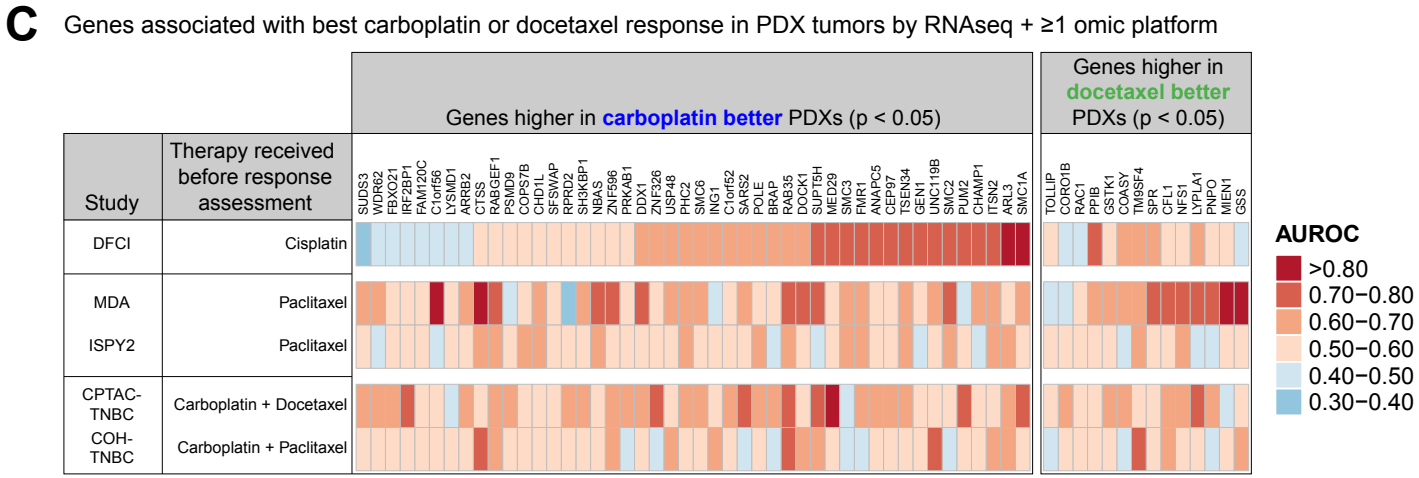
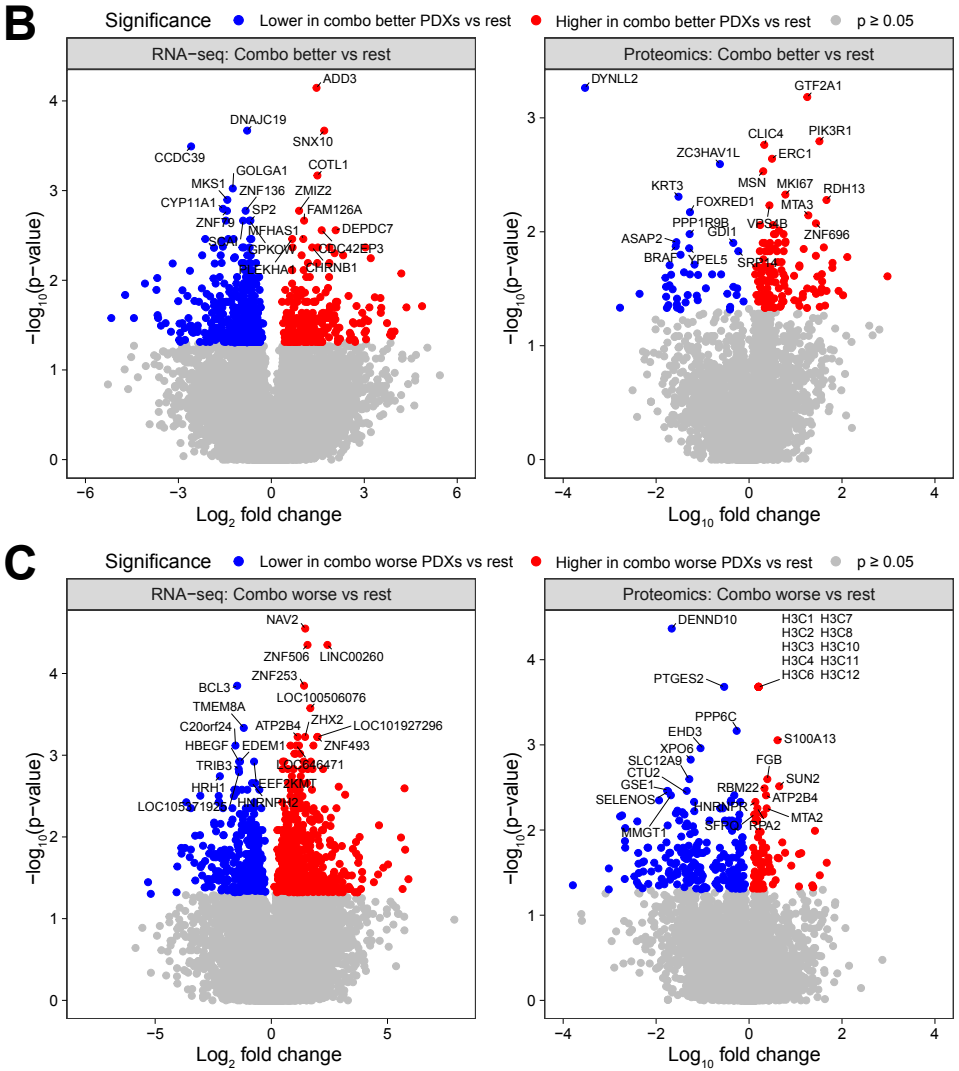


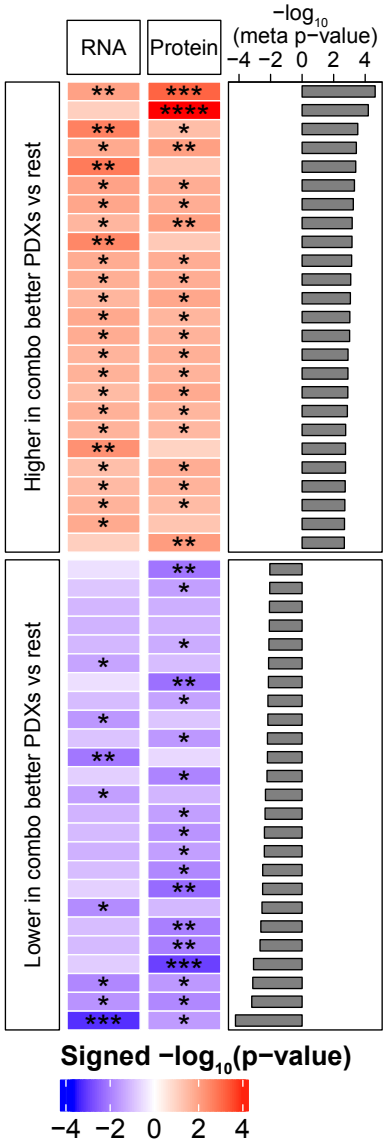
Figure S5: Positive vs negative interactions between agents

A Statistical approach

Wilcoxon-rank sum test:

1. Combo better than single agents (N=4) vs rest (N=38)
2. Combo worse than best single agent (N=5) vs rest (N=37)





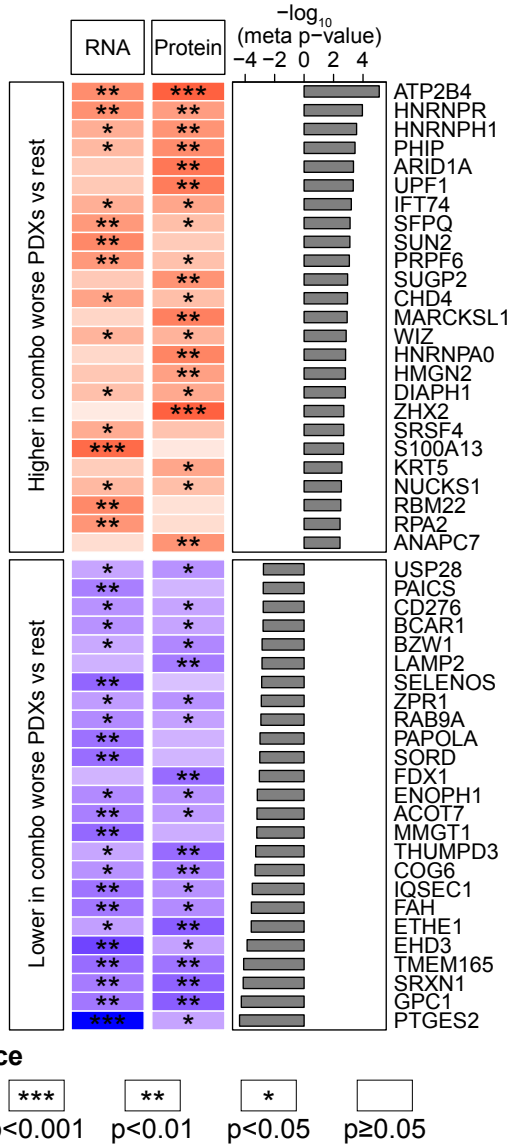
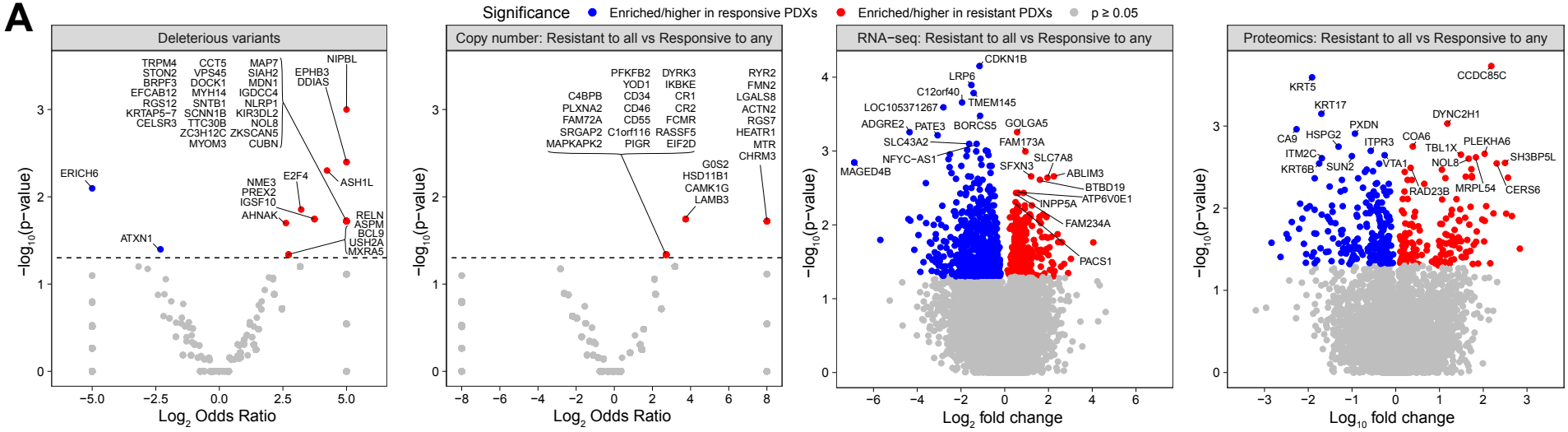
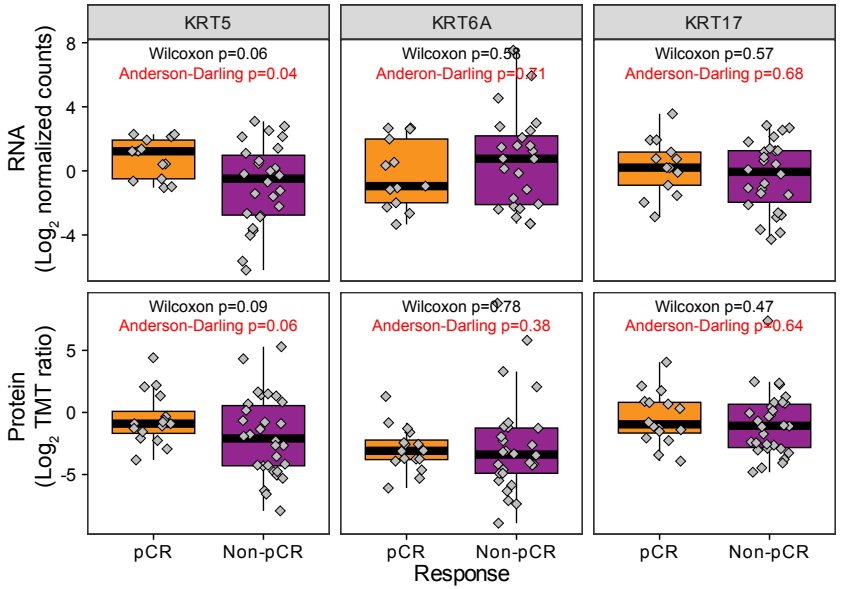
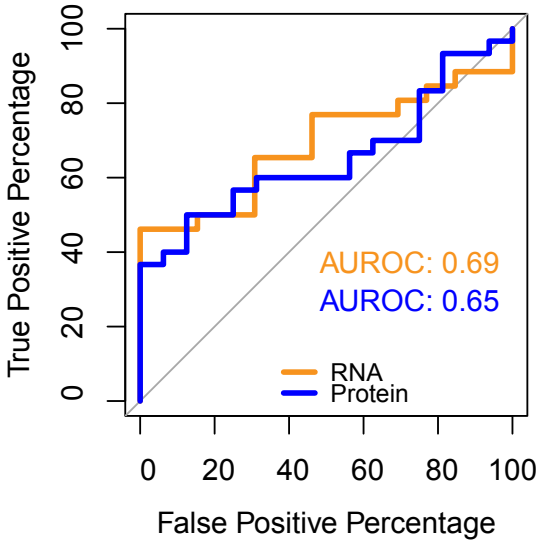


Figure S6: Associations with lack of chemotherapy response



B Cytokeratin association in CPTAC-TNBC dataset



C KRT5 discrimination of non-pCR in CPTAC-TNBC dataset