**RESEARCH**

# Transcription factor network analysis based on single cell RNA-seq identifies that Trichostatin-a reverses docetaxel resistance in prostate Cancer

Patricia M. Schnepp[1], Aqila Ahmed[1], June Escara-Wilke[1], Jinlu Dai[1], Greg Shelley[1], Jill Keller[1,2], Atsushi Mizokami[3] and Evan T. Keller[1,2,4,5*]

## Abstract

**Background:** Overcoming drug resistance is critical for increasing the survival rate of prostate cancer (PCa). Docetaxel is the first cytotoxic chemotherapeutical approved for treatment of PCa. However, 99% of PCa patients will develop resistance to docetaxel within 3 years. Understanding how resistance arises is important to increasing PCa survival.

**Methods:** In this study, we modeled docetaxel resistance using two PCa cell lines: DU145 and PC3. Using the Passing Attributes between Networks for Data Assimilation (PANDA) method to model transcription factor (TF) activity networks in both sensitive and resistant variants of the two cell lines. We identified edges and nodes shared by both PCa cell lines that composed a shared TF network that modeled changes which occur during acquisition of docetaxel resistance in PCa. We subjected the shared TF network to connectivity map analysis (CMAP) to identify potential drugs that could disrupt the resistant networks. We validated the candidate drug in combination with docetaxel to treat docetaxel-resistant PCa in both in vitro and in vivo models.

**Results:** In the final shared TF network, 10 TF nodes were identified as the main nodes for the development of docetaxel resistance. CMAP analysis of the shared TF network identified trichostatin A (TSA) as a candidate adjuvant to reverse docetaxel resistance. In cell lines, the addition of TSA to docetaxel enhanced cytotoxicity of docetaxel resistant PCa cells with an associated reduction of the IC50 of docetaxel on the resistant cells. In the PCa mouse model, combination of TSA and docetaxel reduced tumor growth and final weight greater than either drug alone or vehicle.

**Conclusions:** We identified a shared TF activity network that drives docetaxel resistance in PCa. We also demonstrated a novel combination therapy to overcome this resistance. This study highlights the usage of novel application of single cell RNA-sequencing and subsequent network analyses that can reveal novel insights which have the potential to improve clinical outcomes.

**Keywords:** Docetaxel resistance prostate cancer, Trichostatin a, Single cell RNA sequencing, PANDA method, Transcription factor network analysis

*Correspondence: etkeller@med.umich.edu
[5] Single Cell Spatial Analysis Program, University of Michigan, NCRC B14 RM116, Ann Arbor, MI 48109, USA
Full list of author information is available at the end of the article

Schnepp *et al. BMC Cancer* (2021) 21:1316

Page 2 of 14

## Background

Prostate cancer (PCa) is the second leading cause of cancer-related deaths in men in the United States [1]. While the first line of treatment for advanced PCa is androgen deprivation therapy, the majority of patients develop castrate-resistant PCa (CRPC) [2] which leads to use of chemotherapy. Docetaxel, a taxane, was one of the first cytotoxic therapies approved for CRPC in the United States [3]. It operates through the stabilization of microtubules and inhibition of Bcl-2 expression [4–6]. However, the survival benefits of docetaxel are limited with resistance developing in nearly 99% of patients within 3 years [7]. Understanding how this resistance arises is critical to identify strategies to overcome resistance and increase the survival of PCa patients.

While previous studies to delineate the mechanisms of docetaxel resistance in PCa have identified putative targets, these studies focused on a small number of gene expression changes that occur during drug resistance [8, 9]. In a previous study, we used single cell RNA-sequencing of docetaxel-resistant PCa cells to identify putative candidates of docetaxel resistance [10]. However, a limitation of that study was the lack of integrating the data with gene pathways and transcriptional activators in a more holistic fashion. An integrative and systems-level approach that, in addition to transcription expression, incorporates protein interactions and transcriptional activation may help to better understand the progression towards drug resistance and identify combination therapies to overcome resistance. The ability of such multifaceted integrated approaches that include information from multiple data sources to unveil important biological insights have become apparent in recent years [11–14]. In the current study, we adapted an integrative network inference method, Passing Attributes between Networks for Data Assimilation (PANDA), to model the transcription factor (TF) regulatory network in docetaxel sensitive and resistant PCa cell lines [15]. PANDA develops a regulatory model by iteratively integrating the information from TF-TF protein interaction, gene expression profile and gene co-regulation, and TF-binding motif data. This method has been previously successfully adapted to study ovarian cancer [16] and breast cancer [17] through analysis of bulk samples. In the current study, we applied PANDA for the first time to single cell transcriptomes. Single cell RNA sequencing (scRNA-seq) uncovers the variability and heterogeneity of individual cells in a population that cannot be appreciated using traditional bulk sequencing. This allows us to identify new information only observed through sequencing individual cells and provides a novel application of PANDA method to identify active networks in the development of docetaxel resistance.

In this study, we applied PANDA to characterize the TF regulatory network underlying development of docetaxel resistance in docetaxel sensitive and resistant variants of the PC-3 and Du145 PCa cell lines. We conducted scRNA-seq on the sensitive and resistant variations of both cell lines. We identified shared network nodes and edges between the two cell lines. We also identified the TFs driving resistance and validated their importance in maintaining drug resistance. Furthermore, we subjected the networks to connectivity map analysis (CMAP) to identify candidate therapeutics to reverse docetaxel resistance in PCa. Based on the CMAP, we identified trichostatin A as a candidate therapy and then demonstrated that TSA, in combination with docetaxel successfully decreased tumor growth in both in vitro and in vivo PCa models. This work provides valuable insight into a novel strategy using scRNA-seq to identify mechanisms of docetaxel resistance as well as candidate therapies to reverse drug resistance.

## Materials and methods

### Cell lines and reagents

DU145 (cat no. HTB-81) and PC3 (cat no. CRL-1435) were purchased from ATCC (Virginia, USA). The docetaxel resistant strains were created as previously described [9]. All cells were cultured in RPMI 1640 (Invitrogen Co., Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Life Technologies, Inc.). Resistance was maintained in the cells using growth media supplemented with 10 nM of docetaxel while sensitive cells were maintained with the addition of DMSO to a final level of 0.1% in the growth media (Cell Signaling Technology). Cell identification is confirmed annually using PCR for short tandem repeats.

### Gene expression quantification

The single cell samples were previously sequenced and published by our group [10]. In brief, for 1 week, cells were transferred to docetaxel free media. Cells were trypsinized in 0.05% Trypsin EDTA for 5–10 min at 37 °C and washed with media. For single cell sequencing, the cell suspension was loaded into in the Fluidigm C1™ machine and processed into single cell cDNA libraries according to manufacturer protocol (PN 101–4981). Briefly, full length mRNA-seq libraries were generated from single cells captured using the Fluidigm C1™ Single Cell mRNA Seq IFC, 10-17 μm (PN 100–5760) and Fluidigm C1™ Single-Cell Reagent Kit for mRNA Seq (PN 100–6201). Each chip was visually inspected to identify which wells contained cells. Wells containing one cell were included in library preparation ad sequencing. The capture rate was between 78 and 96% across all chips used in this study. Full length cDNA was converted into

Schnepp *et al. BMC Cancer*    (2021) 21:1316

Page 3 of 14

sequence ready libraries using SMART-seq v4 Ultra Low Input RNA kit for Fluidigm C1™ System (Takara Bio, Mountain View, CA, Cat 635,025) and SeqAmp™ DNA Polymerase (Takara Bio, Cat 638,504). Library preparation was completed using Nextera XT DNA library prep kit (Illumina, San Diego, CA, Cat. FC-131-1096) and Nextera XT DNA Library Prep Index Kit (Illumina, Cat FC-131-1002). Samples following PCR reactions as called for in each kit's manufacturer's protocol was purified using Agencourt AMPure XP (Beckman Coulter, Brea, CA, Item No A63880). Samples were sequenced on Illumina HiSeq-2500 Rapid for DU145 cell line variants and Illumina HiSeq-4000 with single end option for PC3 cell line variants. Reads that were below the minimum quality controls were discarded. Each sample was aligned to the Human Genome hg38 [18] using bowtie alignment tool [19]. We captured a total of 324 cells across all cell lines. Poor quality cells were removed based on low number of reads as determined using the Fluidigm Singular package (https://www.fluidigm.com/software/). A total of 12 cells were removed. 64 DU145 sensitive cells, 71 DU145 resistant cells, 89 PC3 sensitive cells and 88 PC3 resistant cells were included in all downstream analysis. To identify genes for downstream analysis, we used the Fluidigm Singular package. Genes that were expressed in at less than 10 cells in each cell line were excluded. For the remaining genes, the lowest 15% of expressed genes were excluded. Lastly, genes need to be identified in all four cell line samples to be included in the final gene list. This resulted in 12,862 genes being included for all the subsequence downstream analyses. For gene expression analysis, we followed the Seurat pipeline [20]. In brief, we used this pipeline to conduct dimensional reduction (including PCA and tSNE) on all high-quality single cells using all 12,862 genes. Additionally, we estimated the cell cycle status of each cell using the suggested pipeline for the Seurat package. We used the cell cycle markers included in the Seurat package [21].

## Constructing PANDA regulatory networks

PANDA [15, 22] uses three inputs: a motif prior, a set of known TF-TF interactions, and expression data. To create each cell line specific transcriptional regulatory networks, we ran PANDA with the same TF motif prior data set and TF-TF interaction data, but with gene expression unique to each cell line. To create a motif prior data set, we downloaded the *Homo Sapiens* TF motifs from the Catalog of Inferred Sequencing Binding Preferences CIS-BP [23] for the 240 TFs included in the gene expression data sets. The TF position weight matrices were mapped to the promoter regions of all genes (defined as $[-750:+250]$ around the transcription start site for each gene) using FIMO [24]. To control the TF-TF interaction

data set, we estimated the protein-protein interaction network between all 240 TFs using the interaction scores from StringDb v10.5 [25]. All data sources provided in StringDB were included when determining the initial interaction scores. The interaction scores were divided by 1000 and self-interactions were set equal to one. For each network, we constructed a pairwise co-expression levels between each of the target genes (based on Pearson correlation). PANDA then combined this information with the TF motif prior network and TF-TF interaction network to produce each TF regulatory network.

## Specificity score of edges

We identified the enriched edges as calculated in [22]. In brief: for the specificity score (s) of each edge in the regulatory networks, using all four networks, we first calculated the median and interquartile range (IQR) for each edge weight (w) between each TF (t) and gene (g). Next, we compared each individual edge weight (w) to its median and IQR to get the specificity score. An edge was defined as enriched to a network if $s > N$. $N$ was determined by calculating the specificity scores for the individual genes (*g*) by comparing the median expression of the gene $w_g^{(c)}$ in a particular cell line (*c*) to the median and IQR range of the networks constructed from either gene expression data sets. We then varied $N$ from 0 to 1. We selected the cut-off of $N = 0.4$ for the single cell sequenced cell lines since at those cut-offs half of all genes are identified as network enriched.

## Node enrichment

To select the TF and gene node enrichment, we followed the method presented in [16, 22]. In brief, each TF was determined to connect with the number of enriched edges (as determined above) in each PANDA network. Using a hypergeometric distribution, we determined which networks targeted a higher number of enriched edges in either network. We calculated the edge weight change by calculating the average edge weight connected to each TF and took the difference between the two PANDA networks. We selected "key" TF and gene nodes that had a *p*-value less than 0.05 in both the network comparison of DU145 and network comparison of PC3. TF and gene nodes must also have both a positive or negative edge weight fold change in the comparison of DU145 and PC3 sensitive and resistant networks to be considered a "key" TF or gene node.

## Gene set enrichment analysis

TF specificity score for each gene was determine using the specificity scores for each edge connected to the specific TF. Then Gene Set Enrichment Analysis (GSEA) was performed as previously described [22, 26]. In our study,

Schnepp *et al. BMC Cancer*    (2021) 21:1316

Page 4 of 14

we used the list of specific scores for each TF to run a pre-ranked GSEA [26] to test for enrichment of gene ontology terms. Highly significant enriched associations (FDR < 0.05) from these analyses were subjected to hierarchical clustering to group the enriched gene sets into clusters. For each cluster, the frequency that each word appeared in the GO terms assigned to each cluster was determined and used to calculate its statistical enrichment based on its hypergeometric probability. We then scaled the resulting $p$-values by $-\log_{10}$ so that the most statistically relevant words would appear the largest. We used the R package wordcloud2 to generate the resulting word clouds for each cluster.

### Connectivity map analysis

We used the Connectivity Map data set (build01) containing genome-wide expression data for 453 treatment and vehicle control pairs, representing 164 distinct small molecules (https://portals.broadinstitute.org/cmap) [27]. Genes were marked as up- or down-regulated based on their average edge weight in the final TF network.

### Drug response

DU145 and/or PC3 docetaxel-resistant cells were plated and cultured in 96-well plates for 16 h ($2 \times 10^3$ cells/well). Cells were treated with TSA (15 nM; Sigma Aldrich, St. Louis, MO), kaempferol (0.2 μM; Sigma Aldrich), vorinostat (0.15 μM; Tocris, Avonmouth, Bristol) and/or docetaxel (10 nM, Cell Signaling Technology, Danvers, MA) or vehicle (PBS) for 48 h at which time WST-1 solution (Roche Applied Science) was added to the culture medium and incubated for 2.5 h at 37 °C. Absorbance was subsequently determined on a plate reader at a wavelength of 490 nm (Multi-Mode Microplate Reader, SpectraMax M5, Molecular Devices MDS Analytical Technologies).

### Prostate cancer cell growth assay

The cells (3000/well) were seeded in 96-well plates (Corning, New York, NY, USA) in triplicates for 24 h and then the cells were treated with different concentrations of the indicated Trichostatin A (Sigma Chemical Co., St. Louis, MO). The cells were cultured for 3 days. The number of viable cells was measured by Cell Proliferation Reagent WST-1 (Roche Applied Science) as directed by the manufacturer. IC50, the concentration that caused 50% inhibition of cell growth, was calculated with software from AAT Bioquest, Inc. (https://www.aatbio.com/tools/ic50-calculator).

### Mouse experiments

The mouse experiments were approved by the University of Michigan Institutional Animal Care & Use Committee under Protocol 10,366. Docetaxel-resistant PC3 cells in PBS + 50% matrigel GFR (Corning, Corning, New York) were injected subcutaneously into 60 SCID mice (male, 6 to 9 weeks old, mice sourced from the Unit for Laboratory Animal Mice Breeding Colony Managers at University of Michigan (ULAM-BCM) with approval from them for our experiments) at $10^6$ cells per right flank. During experiment, when specified in the approved protocol, mice were anesthetized with isoflurane inhalation at 2.5 mg/kg. Tumor growth was monitored using calipers to measure the length (L) and width (W) and tumor volume was calculated using the formula π/6 x $W^2$ x L. On Day 15 post-injection when average tumor volume was 100mm$^3$, mice were randomly assigned to one of 4 treatment groups: vehicle 5% DMSO ($n = 13$), 5 mg/kg docetaxel ($n = 13$), 1.5 mg/kg TSA (Sigma, St. Louis, MO), or combination 5 mg/kg docetaxel and 1.5 mg/kg TSA treatment ($n = 13$). Docetaxel was injected intraperitoneally (IP) weekly in the docetaxel alone and combination groups. TSA was injected IP three times per week on Monday, Wednesday and Friday to the TSA alone and TSA-docetaxel Combo group. At the end of experiment, all mice were euthanized with $CO_2$ inhalation and cervical dislocation as secondary method.

### Statistical analysis

For animal experiments, power calculations were performed using Gpower 3.1.9.7 [28]. For a power of 90% and an effect size of 55% with $p < 0.05$, 13 animals per group were required. Statistical comparison among groups were calculated using a mixed-effects analysis for tumor growth over time and a one-way ANOVA followed by Tukey's Honest Significant Difference for difference in tumor weight. Statistical significance was determined as $p \leq 0.05$.

## Results

### Identifying network enriched edges

We initially conducted single cell RNA-sequencing on two PCa cell lines, DU145 and PC3, that had been previously created to be resistant to docetaxel treatment [9]. We sequenced both the parental (i.e. docetaxel-sensitive) and docetaxel-resistant variants from both cell lines resulting in a total 312 individual sequenced cells (64 DU145 sensitive cells, 71 DU145 resistant cells, 89 PC3 sensitive cells and 88 PC3 resistant cells). To construct TF activity networks for each of the four established cell lines, we used the PANDA algorithm [15]. PANDA integrates the gene-gene co-expression information from each of the four established cell lines with an initial regulatory network consisting of 240 TF as well as known TF-TF interactions and TF-Gene interactions [25]. This resulted in 4 reconstructed TF regulatory

Schnepp *et al. BMC Cancer*    (2021) 21:1316

Page 5 of 14

networks (Fig. 1A). We visualized the heterogeneity of all 312 cells and observed that the majority of cells clustered based primarily on their cell line identity (Fig. 1B and Supplementary Fig. S1). We did see a change in the cell cycle phase between the sensitive and resistant cell lines (Supplementary Table S1). We observed an over

two-fold decrease in S and G2M phase cells in the resistant compared to sensitive cells from both lines. However, the G1 phase cells increased by over two and half fold in the resistant cells compared to the sensitive cells (Supplementary Table S1). This does suggest a shift in cell cycle stage following docetaxel resistance in PCa cells. To



**Fig. 1** Detailed Workflow of Study. **A** Gene expression data from each scRNA-seq of each cell was combined with physical protein-protein interactions and predicted TF-gene targets to build individual network models. The models from each cell line were compared to identify differences. Shared differences between the DU145 models and PC3 models were combined to create a general model of docetaxel drug resistance. For representation: circles denote TF, squares denote a gene. In the combined network: grey denotes the node or edge is not significantly altered between the sensitive and resistant cells, a green edge or node is statistically significant. **B** tSNE of all single cells analyzed

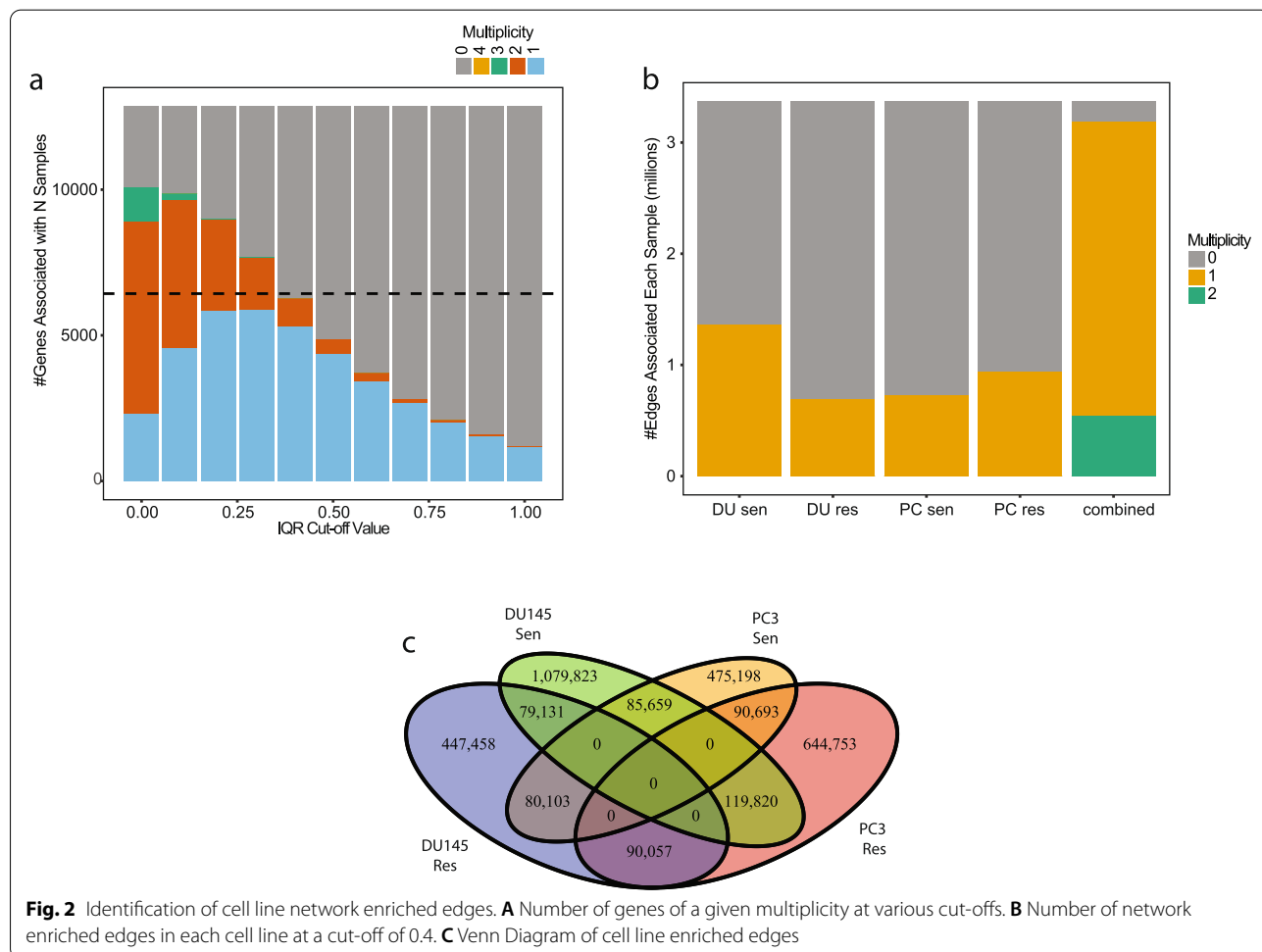Schnepp *et al. BMC Cancer*     (2021) 21:1316

Page 6 of 14

determine which edges and nodes were enriched in the sensitive and resistant DU145 or PC3 networks, we first calculated the interquartile range (IQR) cut-off in which half of the nodes (based on gene expression) would be labeled as network enriched as previously described [22]. We determined this value to be a cut-off of 0.4 (Fig. 2A-B). 3,022,871 edges (89.7%) were enriched in at least one network (Fig. 2B). Of those, only 375,639 edges were enriched for two networks (Fig. 2C). The observation that only 11.1% of the edges overlap between any two networks, indicates that these edges primarily illustrate cell line enriched regulatory interactions across each condition.
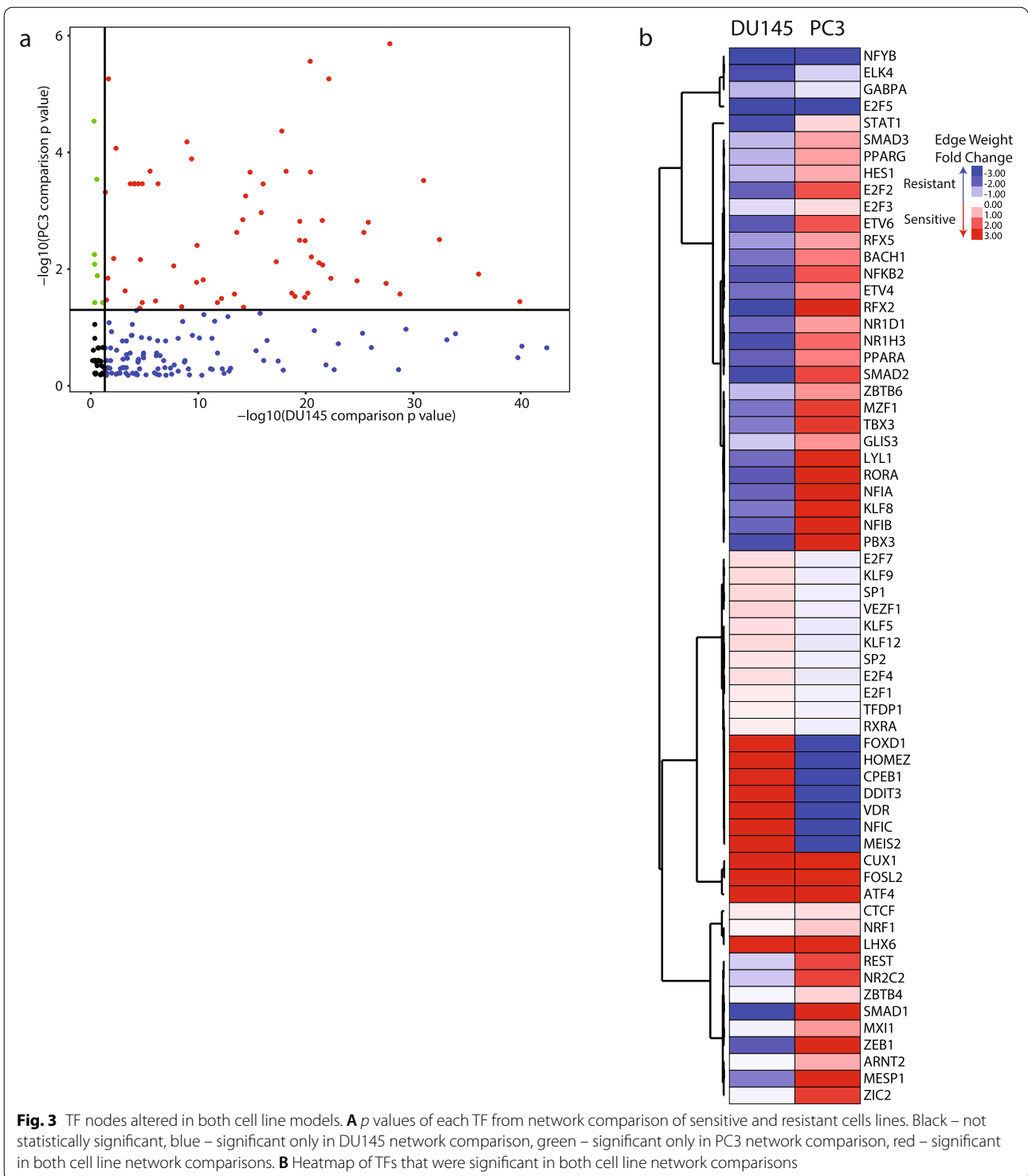
### Identifying cell line specific nodes

To determine which TF and gene nodes differed between the sensitive and resistant DU145 or PC3 networks, we compared the 'in-degree' of each node as defined as the sum of all edge weights connected to a particular node for each network. We calculated the probability for node statistical significance between the two DU145 or PC3 networks by comparing the number of enriched edges targeted by each node in either network [16]. Of the 240 included TFs, 63 had a *p*-value < 0.05 in both cell line network comparisons (Fig. 3A and Supplementary Table S2). Additionally, we calculated the edge weight fold change based on the 'in degree' value for each node. For the 63 TFs, not all had the same fold change direction change for both comparisons. Only 10 of the TFs had an in-degree in the same direction (Fig. 3B).

We also determined which gene nodes were altered in the TF regulatory networks using the 'in degree' values calculated for each gene node. Of the 12,862 genes included in the original gene expression data sets, 210 gene nodes had a *p*-value < 0.05 in both cell line network comparisons (Fig. 4A and Supplementary Table S3). 118 gene nodes had a fold change direction that was the same between the DU145 and PC3 network comparisons (Fig. 4B).



**Fig. 2** Identification of cell line network enriched edges. **A** Number of genes of a given multiplicity at various cut-offs. **B** Number of network enriched edges in each cell line at a cut-off of 0.4. **C** Venn Diagram of cell line enriched edges

Schnepp *et al. BMC Cancer*      (2021) 21:1316

Page 7 of 14



**Fig. 3** TF nodes altered in both cell line models. **A** p values of each TF from network comparison of sensitive and resistant cells lines. Black – not statistically significant, blue – significant only in DU145 network comparison, green – significant only in PC3 network comparison, red – significant in both cell line network comparisons. **B** Heatmap of TFs that were significant in both cell line network comparisons
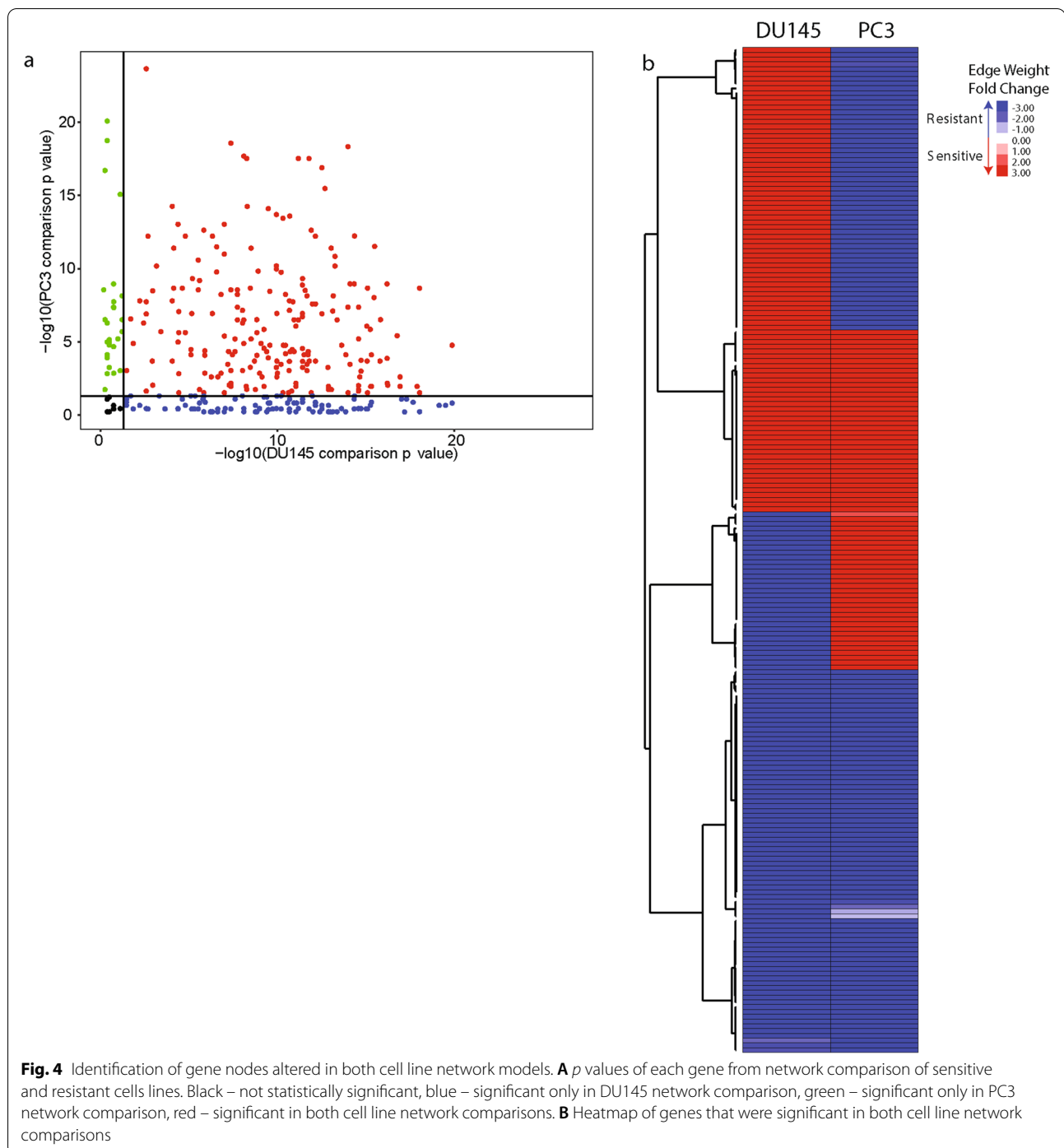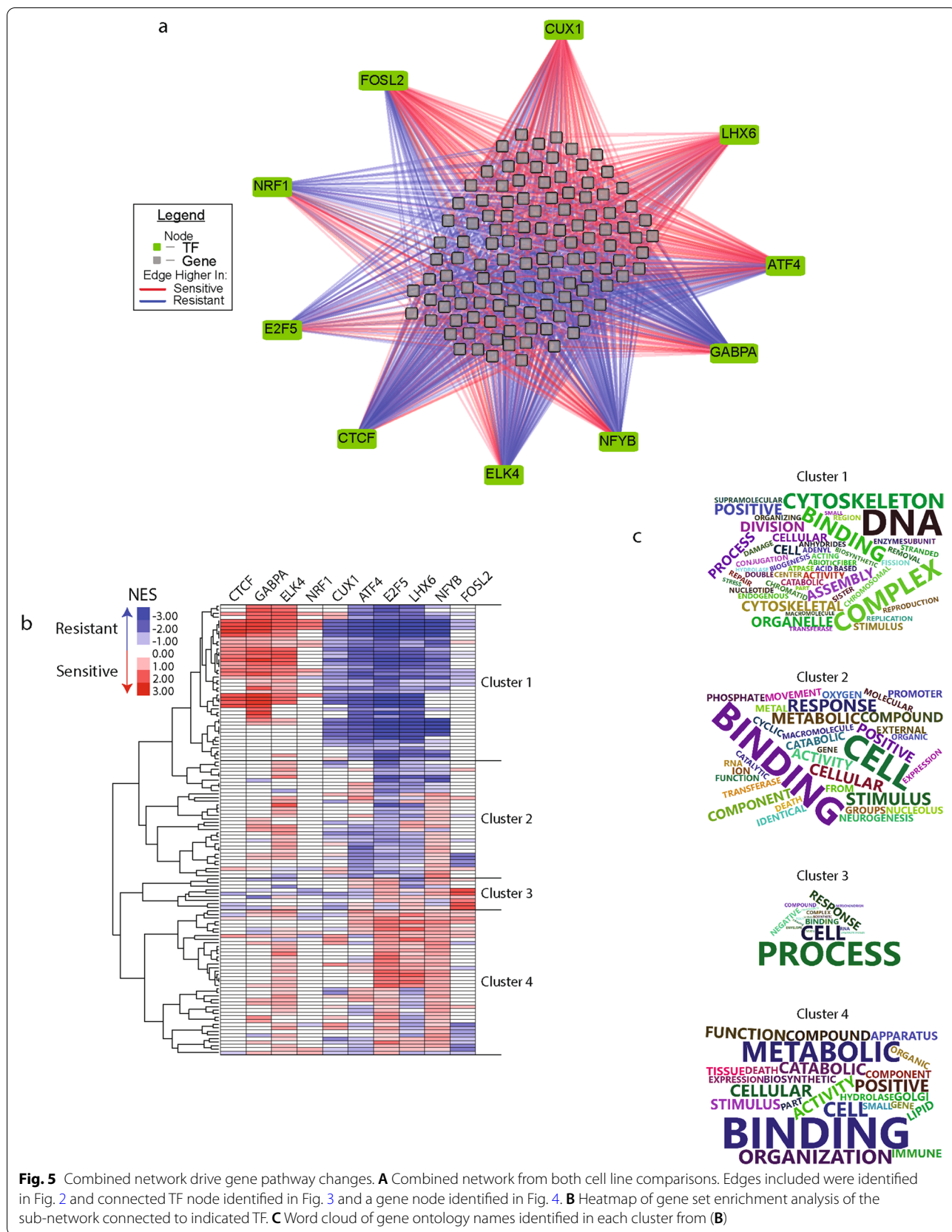
## Regulation of functional pathways

To ensure the final network was not representative of just one PCa cell line and to provide robust results, we constructed a final combined TF activity network. This network only includes the common edges and nodes

identified in Figs. 2, 3 and 4. Thus, the final TF activity network represents a generalized prostate cancer response to docetaxel treatment (Fig. 5A). In this visualization, the 10 TFs were connected to the 118 gene nodes by lines colored based on whether they exhibit higher

**Fig. 4** Identification of gene nodes altered in both cell line network models. **A** *p* values of each gene from network comparison of sensitive and resistant cells lines. Black – not statistically significant, blue – significant only in DU145 network comparison, green – significant only in PC3 network comparison, red – significant in both cell line network comparisons. **B** Heatmap of genes that were significant in both cell line network comparisons

average edge weight in the sensitive (red) or resistant (blue) regulatory networks. To explore the functional pathways altered in the regulatory network, we ran gene set enrichment analysis (GSEA) on the TF specific targeted genes to identify four clusters of groups of GO terms (Fig. 5B and Supplementary Table S4). We used word clouds to summarize the GO terms for each cluster

to provide a snapshot of the functions (Fig. 5C). These clusters often include sets of highly related functions. Cluster 1 contains GO terms related to the cytoskeleton, chromatin and cellular division suggesting these terms regulate cellular proliferation. In cluster 2, we observe GO terms related to metal ions, signaling molecules and binding suggesting these GO terms regulate cellular

Schnepp *et al. BMC Cancer*    (2021) 21:1316

Page 9 of 14



**Fig. 5** Combined network drive gene pathway changes. **A** Combined network from both cell line comparisons. Edges included were identified in Fig. 2 and connected TF node identified in Fig. 3 and a gene node identified in Fig. 4. **B** Heatmap of gene set enrichment analysis of the sub-network connected to indicated TF. **C** Word cloud of gene ontology names identified in each cluster from (**B**)

Schnepp *et al. BMC Cancer*    (2021) 21:1316

Page 10 of 14

signaling. Cluster 3 is the smallest of all the clusters and contains words such as cellular process, envelope, and RNA. This is suggesting this cluster regulates cellular response to signaling but not through any specific mechanism like in cluster 2. Cluster 4 contains words such as metabolic and catabolic suggesting this cluster contains GO terms relating to cellular metabolism Additionally, cluster 1 is enriched by different TFs in the resistant or sensitive cell lines. This may suggest the importance of these pathways regardless of cellular drug sensitivity. In cluster 4, the majority of enrichment was higher in the sensitive cell lines when targeted by the TFs. These data suggest the functional pathway changes that occur during docetaxel resistance due to changes in TF activity.

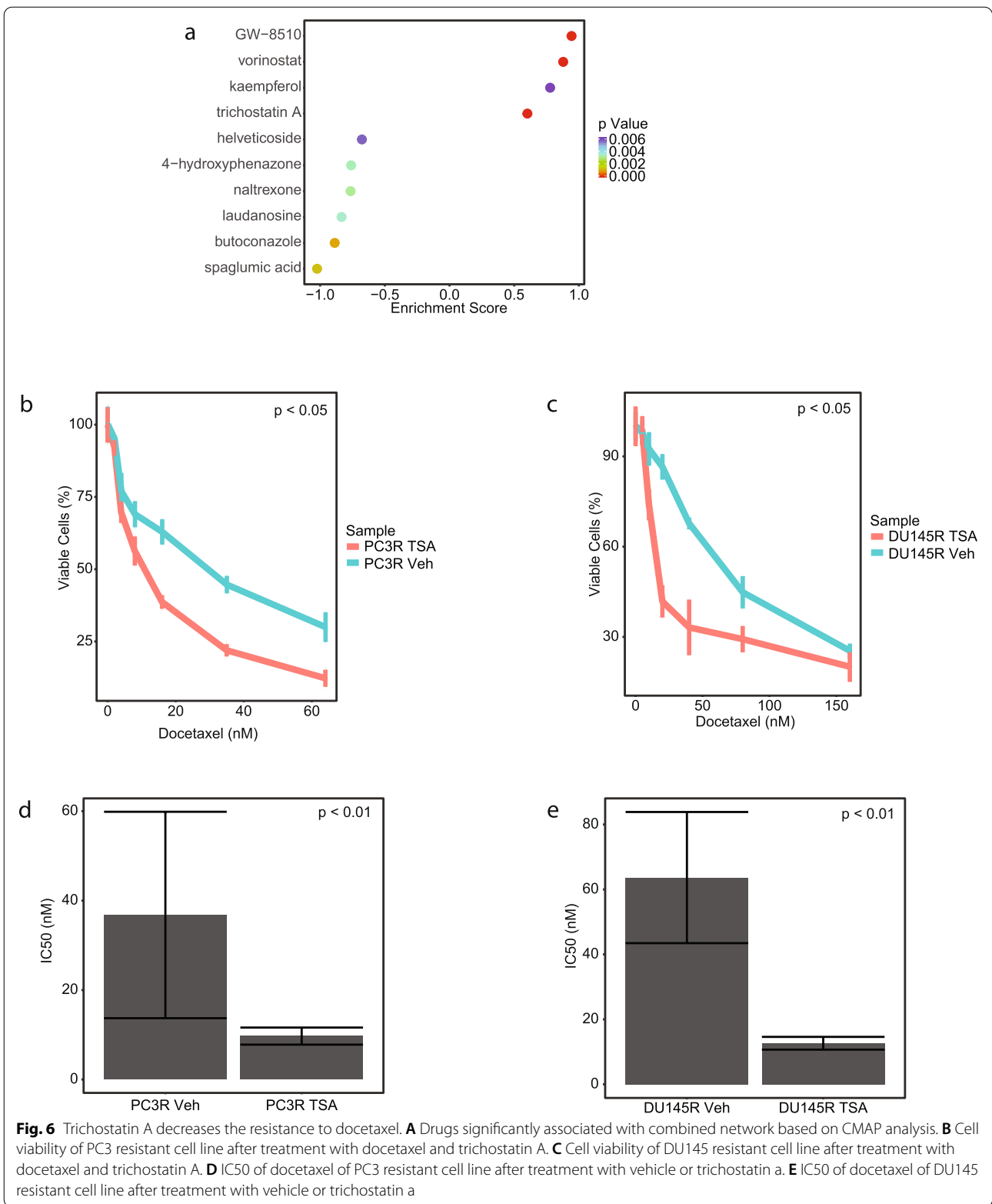### Drug combination therapy for drug resistant PCa

To identify potential alternative clinical therapeutic opportunities in our PCa TF network, we investigated drugs that would potentially disrupt the combined TF regulatory network. To perform this, we labeled the 118 gene nodes as sensitive or resistant nodes based on their edge weight fold change (Fig. 2B). Using this list, we used Connectivity Map (CMAP) analysis to predict drugs that would up-regulate the sensitive genes nodes and down-regulate the resistant gene nodes. The drugs with positive enrichment would be those with the highest potential to reverse docetaxel resistance in PCa. We examined the top hits from CMAP using the final combined regulatory network and identified four potential drugs: vorinostat, GW-8510, kaempferol, and trichostatin A (Fig. 6A and Supplementary Table S5). To investigate these drugs, we tested their ability to overcome docetaxel resistance in the two PCa cell lines. However, we could not evaluate GW-8510 as we were not able to procure it. After determining the IC50 dosage for the remaining three drugs (Supplementary Fig. S2), we found that neither kaempferol nor vorinostat had a significant impact on docetaxel resistance in the PCa cells (Supplementary Fig. S3). However, the combination of trichostatin A and docetaxel statistically significantly decreased the cellular proliferation of both cell lines compared to the vehicle (Fig. 6B-C). Additionally, the IC50 of docetaxel decreased significantly when the same cells were treated with trichostatin A indicating that TSA decreased resistance to docetaxel (Fig. 6D-E). We further tested the combination treatment of trichostatin A and docetaxel in a mouse model of PCa. We established subcutaneous tumors of PC3 resistant cells in mice and then treated the mice with vehicle, docetaxel alone, TSA alone, or the combination of docetaxel and TSA for 13 days. While neither drug alone impacted tumor growth, the combination of TSA and docetaxel decreased tumor growth as measured by both tumor volume (Fig. 7A, *p*-value < 0.05) and final tumor
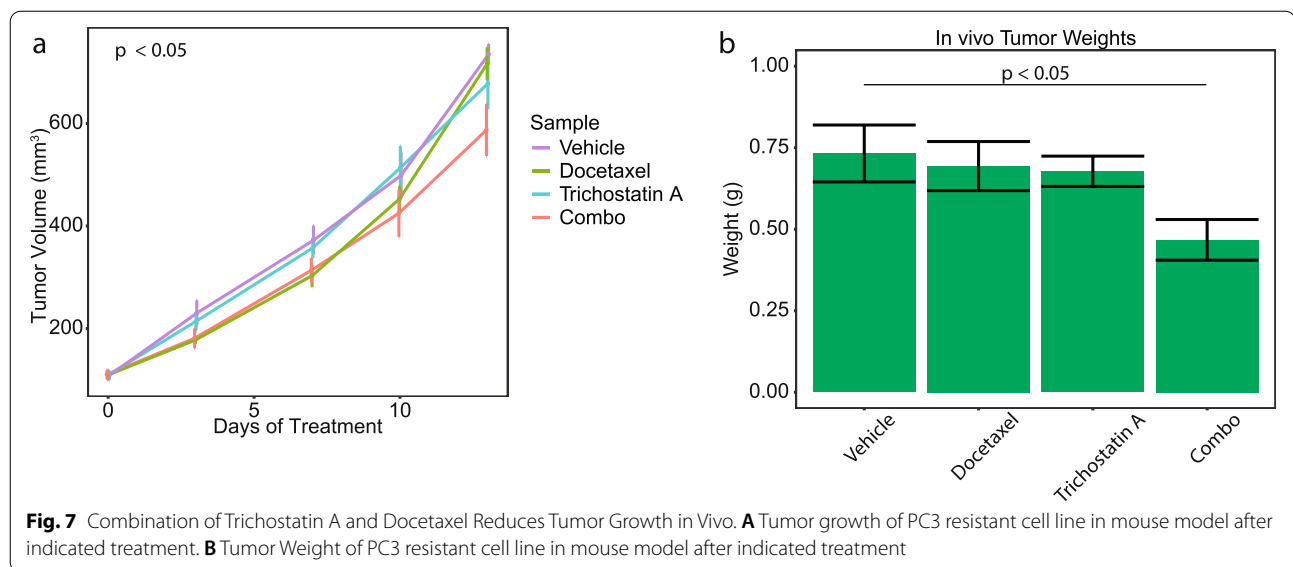
weight (Fig. 7B, *p*-value < 0.05) compared to the three other groups. These data suggest trichostatin A is able to reverse docetaxel resistance in PCa cells in both in vitro and in vivo models.

### Discussion

In this study, we conducted a single cell-based TF network analysis of docetaxel resistance in PCa. We modeled the TF regulatory networks in docetaxel sensitive and resistant PC3 and DU145 PCa cell lines. The final network identified 10 TFs that were the main nodes for the regulatory network. This suggests these TFs are critical for the development of docetaxel resistance in PCa. Furthermore, the gene nodes from the network analysis were subjected to CMAP which suggested that TSA could reverse docetaxel resistant. These findings were validated through use of both in vitro and in vivo models that demonstrated TSA reverses docetaxel resistance. In cell lines, the combination of TSA and docetaxel significantly reduced both the number of viable docetaxel resistant PCa cells and the IC50 value for those resistant cells (Fig. 6). In the PCa mouse model, the combination of TSA and docetaxel reduced tumor growth and final tumor weight greater than either drug alone (Fig. 7). Taken together, these finding demonstrate the validity of this novel methodology of applying network analysis to single cell transcriptomic data to analyze mechanisms of therapeutic resistance and highlights a specific drug that can be tested as a candidate to overcome docetaxel resistance in patients.

The combination of both scRNA-seq and network analysis has enabled the investigation into the TF activity underlying PCa docetaxel resistance. TF activity is regulated beyond the expression level through post-translational modifications including acetylation [29], ubiquitination [30], and sumoylation [31] among others. The application of network analysis, allows for investigation of TF activity beyond just TF gene expression. PANDA of bulk mRNA was previously used to identify the TF co-regulation that existed in triple-negative breast cancer [17] as well as potential biomarkers for anti-angiogenesis treatment in ovarian cancer [16]. In our study, PANDA allowed for determining TF activity networks that identified novel targets for PCa resistance. Our analysis also identified different patterns of TF targeted pathway activation in our GSEA analysis and in the word cloud representation. The word cloud representation identified very high-level terms suggesting that the development of docetaxel resistance impacts very basic cell function. For instance, in cluster 1, a group of gene pathways involved in cytoskeleton and cellular division, we observe one set of TFs driving those pathways. However, in the resistant network, a different set of TFs drive

Schnepp *et al. BMC Cancer*     (2021) 21:1316

Page 11 of 14



**Fig. 6** Trichostatin A decreases the resistance to docetaxel. **A** Drugs significantly associated with combined network based on CMAP analysis. **B** Cell viability of PC3 resistant cell line after treatment with docetaxel and trichostatin A. **C** Cell viability of DU145 resistant cell line after treatment with docetaxel and trichostatin A. **D** IC50 of docetaxel of PC3 resistant cell line after treatment with vehicle or trichostatin a. **E** IC50 of docetaxel of DU145 resistant cell line after treatment with vehicle or trichostatin a

Schnepp *et al. BMC Cancer*     (2021) 21:1316

Page 12 of 14



**Fig. 7** Combination of Trichostatin A and Docetaxel Reduces Tumor Growth in Vivo. **A** Tumor growth of PC3 resistant cell line in mouse model after indicated treatment. **B** Tumor Weight of PC3 resistant cell line in mouse model after indicated treatment

these pathways. This suggests a TF shift from CTCF, GABPA and ELK4 to AFT4, E2F5, and LHX6 among other TFs during drug resistance. Additional study would be needed to explore this TF activity shift.

In our analysis, 10 TFs were identified to be of some statistical significance to either the sensitive or resistant networks. Some of these TFs have been implicated in tumor development or drug resistance in previous studies. CUX1, identified in the sensitive network, is a tumor suppressor and it's lose promotes tumorigenesis [32]. In PCa, the loss of CUX1 reduces the level of cellular senescence in tumor cells [33]. Combined with our data that suggests a loss of CUX1 activity in resistance cells, CUX1 may play a role in docetaxel resistance in PCa. However, there are multiple TFs with higher activity in the resistant networks such as GABPA, NFYB, and NRF1 among others. GABPA is a downstream target of the androgen receptor in PCa and enables the tumor cells to become more aggressive [34]. NFYB drives paclitaxel resistance in breast cancer [35] and oxalipatin resistance in colorectal cancer [36]. NRF family can been observed to drive cisplatin resistance in pancreatic cancer [37, 38]. Together, this provides evidence that these TFs can drive drug resistance or aggressiveness in tumor cells. And in conjunction with our analyses, could play a role in docetaxel resistance in PCa. However, additional research would be needed to confirm this role.

Applying CMAP to our PANDA network analyses enabled identification of candidate drugs, such as TSA. TSA is a reversible histone deacetylase inhibitor [39] that has been previously shown to be a promising new treatment for other cancers. In osteosarcoma cells, TSA induced cancer cell apoptosis through both histone acetylation- and mitochondria-dependent mechanisms [40]. Interestingly, TSA had a greater specificity in affecting cancer cells compared to normal cells than other histone deacetylase inhibitors [41]. This procancer selectivity makes TSA an attractive therapeutic agent in a clinical setting. Furthermore, TSA enhanced the anti-tumor effects of docetaxel in lung cancer [42]. It was demonstrated that that TSA in combination with docetaxel reduced lung cancer cells by promoting apoptosis. While further work is necessary to determine if a similar mechanism was involved with TSA and docetaxel in PCa, our findings suggest that TSA could overcome docetaxel resistance in PCa cells in patients.

Next generation sequencing is currently used in the clinic to aid in determining potential therapies, such as the identification of mutations or gene fusions, in a precision medicine approach [43, 44]. However, these studies do not identify intra-patient heterogeneity, which plays a role in a patient's therapeutic response [45, 46]. The addition of single cell sequencing allows for the study of heterogeneity in each patient which could improve a precision medicine approach. Using the analysis pipeline presented in this study, it is possible to identify TF drivers and off-target drug treatments that disrupt individual sub-populations of cells in an individual patient. Further research is needed to determine which sub-populations drive patient treatment responses in order to target the proper sub-populations. Ultimately, this will allow for a personalized therapeutic approach for patients that do not respond to or become resistant to conventional therapies.

Schnepp *et al. BMC Cancer*    (2021) 21:1316

Page 13 of 14

## Conclusion

Overcoming drug resistance is critical to improving patient survival in PCa. In this study, we identified a TF activity network common to two different PCa cell lines that drives docetaxel resistance in PCa. We also demonstrated a novel combination therapy to overcome this resistance. This study highlights the usage of novel application of single cell RNA-sequencing and subsequent network analyses that can reveal novel insights which have the potential to improve clinical outcomes.

## Abbreviations

PCa: Prostate cancer; TF: Transcription factor; CMAP: Connectivity map; TSA: Trichostatin A; PANDA: Passing Attributes between Networks for Data Assimilation; CRPC: Castration-resistant prostate cancer; scRNA-seq: Single cell RNA-sequencing; GSEA: Gene Set Enrichment Analysis; FDR: False discovery rate; GO: Gene ontology; IQR: Interquartile range; ULAM-BCM: Unit for Laboratory Animal Medicine Breeding Colony Managers.

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-021-09048-0.

---

**Additional file 1: Supplementary Fig. 1**. PCA plot of all sequenced single PCa cells. **Supplementary Fig. 2**. Dosage Curves for potential treatment of PCa cells. (A) Dosage curve for 1 trichostatin A. (B) Dosage curve for vorinostat. (C) Dosage curve for kampferol. 2. **Supplementary Fig. 3**. Combination Treatment of PCa Cells with Potential Drugs and Docetaxel. (A) 3 Proliferation of DU145 resistant cells after treatment of docetaxel and kaempferol. (B) Proliferation of PC3 4 resistant cells after treatment of docetaxel and kaempferol. (C) Proliferation of DU145 resistant cells after 5 treatment of docetaxel and vorinostat. (D) Proliferation of PC3 resistant cells after treatment of docetaxel 6 and vorinostat. *: *p* value < 0.005.

**Additional file 2: Supplementary Table 1**. Cell Cycle Counts Across the Sequenced Single Cells. **Supplementary Table 2**. TF node comparison. **Supplementary Table 3.** Gene node comparison. **Supplementary Table 4**. GSEA Normalized Enrichment Scores. **Supplementary Table 5.** Connectivity Map results

---

## Authors' contributions

PMS performed all bioinformatic data analysis. GS, AM and ETK contributed to sequencing of samples and providing cell lines. AA, JD and PMS conducted and analyzed all drug treatment experiments on cell lines. JE-W, GS, JK and PMS conducted and analyzed mouse related experiments. PMS and ETK wrote the main manuscript text. The authors read and approved the final manuscript.

## Availability of data and materials

The single cell RNA sequencing dataset is available from the Gene Expression Omnibus under GSE140440 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE140440].

## Declarations

### Ethics approval and consent to participate

The mouse experiments were approved by the University of Michigan Institutional Animal Care & Use Committee under Protocol 10366. All experiments were carried out as specified in Protocol 10366. We received written consent from ULAM-BCM to use mice sourced from them in our experiments. The study was carried out in compliance with the ARRIVE guidelines. No human experiments were conducted.

### Consent for publication

All authors have consented to publication of this manuscript.

### Competing interests

The authors declare that they have no competing interests.

### Author details

[1]Department of Urology, University of Michigan Medical School, NCRC B14 RM116, Ann Arbor, MI 48109, USA. [2]Unit for Laboratory Animal Medicine, University of Michigan, NCRC B14 RM116, Ann Arbor, MI 48109, USA. [3]Department of Urology, Kanazawa University, Kanazawa, Japan. [4]Biointerfaces Institute, University of Michigan, NCRC B14 RM116, Ann Arbor, MI 48109, USA. [5]Single Cell Spatial Analysis Program, University of Michigan, NCRC B14 RM116, Ann Arbor, MI 48109, USA.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69:7–34.
2. Marques RB, de Ridder CMA, van Weerden WM. Patient-derived xenograft models of prostate Cancer. In: Hoffman RM, editor. Patient-derived mouse models of Cancer : patient-derived Orthotopic xenografts (PDOX), molecular and translational medicine. Cham: Springer International Publishing; 2017. p. 89–112.
3. Hwang C. Overcoming docetaxel resistance in prostate cancer: a perspective review. Ther Adv Med Oncol. 2012;4:329–40.
4. Cortes JE, Pazdur R. Docetaxel. J Clin Oncol. 1995;13:2643–55.
5. Haldar S, Basu A, Croce CM. Bcl2 is the guardian of microtubule integrity. Cancer Res. 1997;57:229–33.
6. Pienta KJ. Preclinical mechanisms of action of docetaxel and docetaxel combinations in prostate cancer. Semin Oncol. 2001;28:3–7.
7. Petrylak DP, Tangen CM, Hussain MHA, Lara PN, Jones JA, Taplin ME, et al. Docetaxel and estramustine compared with mitoxantrone and prednisone for advanced refractory prostate cancer. N Engl J Med. 2004;351:1513–20.
8. Farah E, Li C, Cheng L, Kong Y, Lanman NA, Pascuzzi PE, et al. NOTCH signaling is activated in and contributes to resistance in enzalutamide-resistant prostate cancer cells. J Biol Chem. 2019. https://doi.org/10.1074/jbc.RA118.006983.
9. Takeda M, Mizokami A, Mamiya K, Li YQ, Zhang J, Keller ET, et al. The establishment of two paclitaxel-resistant prostate cancer cell lines and the mechanisms of paclitaxel resistance with two cell lines. Prostate. 2007;67:955–67.
10. Schnepp PM, Shelley G, Dai J, Wakim N, Jiang H, Mizokami A, et al. Single-cell transcriptomics analysis identifies nuclear protein 1 as a regulator of docetaxel resistance in prostate Cancer cells. Mol Cancer Res. 2020;18:1290–301.
11. Barabási A-L. Network medicine--from obesity to the 'diseasome'. N Engl J Med. 2007;357:404–7.
12. Papin JA, Reed JL, Palsson BO. Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. Trends Biochem Sci. 2004;29:641–7.

Schnepp *et al. BMC Cancer*      (2021) 21:1316

Page 14 of 14

13. Silverman EK, Loscalzo J. Developing new drug treatments in the era of network medicine. Clin Pharmacol Ther. 2013;93:26–8.
14. Silverman EK, Loscalzo J. Network medicine approaches to the genetics of complex diseases. Discov Med. 2012;14:143–52.
15. Glass K, Huttenhower C, Quackenbush J, Yuan G-C. Passing messages between biological networks to refine predicted interactions. PLoS One. 2013;8:e64832.
16. Glass K, Quackenbush J, Spentzos D, Haibe-Kains B, Yuan G-C. A network model for angiogenesis in ovarian cancer. BMC Bioinformatics. 2015;16:115.
17. Min L, Zhang C, Qu L, Huang J, Jiang L, Liu J, et al. Gene regulatory pattern analysis reveals essential role of core transcriptional factors' activation in triple-negative breast cancer. Oncotarget. 2017;8:21938–53.
18. Consortium, I.H.G.S. Initial sequencing and analysis of the human genome. Nature. 2001;409:860–921.
19. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.
20. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol. 2015;33:495–502.
21. Kowalczyk MS, Tirosh I, Heckl D, Rao TN, Dixit A, Haas BJ, et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Res. 2015;25:1860–72.
22. Sonawane AR, Platig J, Fagny M, Chen C-Y, Paulson JN, Lopes-Ramos CM, et al. Understanding Tissue-Specific Gene Regulation. Cell Rep. 2017;21:1077–88.
23. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, et al. Determination and inference of eukaryotic transcription factor sequence specificity. Cell. 2014;158:1431–43.
24. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. Bioinformatics. 2011;27:1017–8.
25. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43:D447–52.
26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. PNAS. 2005;102:15545–50.
27. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006;313:1929–35.
28. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*power 3.1: tests for correlation and regression analyses. Behav Res Methods. 2009;41:1149–60.
29. Soutoglou E, Katrakili N, Talianidis I. Acetylation regulates transcription factor activity at multiple levels. Mol Cell. 2000;5:745–51.
30. Conaway RC, Brower CS, Conaway JW. Emerging roles of ubiquitin in transcription regulation. Science. 2002;296:1254–8.
31. Geiss-Friedlander R, Melchior F. Concepts in sumoylation: a decade on. Nat Rev Mol Cell Biol. 2007;8:947–56.
32. Wong CC, Martincorena I, Rust AG, Rashid M, Alifrangis C, Alexandrov LB, et al. Inactivating CUX1 mutations promote tumorigenesis. Nat Genet. 2014;46:33–8.
33. Revandkar A, Perciato ML, Toso A, Alajati A, Chen J, Gerber H, et al. Inhibition of Notch pathway arrests PTEN-deficient advanced prostate cancer by triggering p27-driven cellular senescence. Nat Commun. 2016;7:13719.
34. Sharma NL, Massie CE, Butter F, Mann M, Bon H, Ramos-Montoya A, et al. The ETS family member GABPα modulates androgen receptor signalling and mediates an aggressive phenotype in prostate cancer. Nucleic Acids Res. 2014;42:6256–69.
35. Bauer JA, Ye F, Marshall CB, Lehmann BD, Pendleton CS, Shyr Y, et al. RNA interference (RNAi) screening approach identifies agents that enhance paclitaxel activity in breast cancer cells. Breast Cancer Res. 2010;12:R41.
36. Fang Z, Gong C, Yu S, Zhou W, Hassan W, Li H, et al. NFYB-induced high expression of E2F1 contributes to oxaliplatin resistance in colorectal cancer via the enhancement of CHK1 signaling. Cancer Lett. 2018;415:58–72.
37. Hong YB, Kang HJ, Kwon SY, Kim HJ, Kwon KY, Cho CH, et al. Nrf2 regulates drug resistance in pancreatic cancer cells. Pancreas. 2010;39:463–72.
38. Karathedath S, Rajamani BM, Musheer Aalam SM, Abraham A, Varatharajan S, Krishnamurthy P, et al. Role of NF-E2 related factor 2 (Nrf2) on

chemotherapy resistance in acute myeloid leukemia (AML) and the effect of pharmacological inhibition of Nrf2. PLoS One. 2017;12:e0177227.
39. Vanhaecke T, Papeleu P, Elaut G, Rogiers V. Trichostatin A-like hydroxamate histone deacetylase inhibitors as therapeutic agents: toxicological point of view. Curr Med Chem. 2004;11:1629–43.
40. Roh MS, Kim CW, Park BS, Kim GC, Jeong JH, Kwon HC, et al. Mechanism of histone deacetylase inhibitor Trichostatin a induced apoptosis in human osteosarcoma cells. Apoptosis. 2004;9:583–9.
41. Chang J, Varghese DS, Gillam MC, Peyton M, Modi B, Schiltz RL, et al. Differential response of cancer cells to HDAC inhibitors trichostatin a and depsipeptide. Br J Cancer. 2012;106:116–25.
42. Zhang Q-C, Jiang S-J, Zhang S, Ma X-B. Histone deacetylase inhibitor trichostatin a enhances anti-tumor effects of docetaxel or erlotinib in A549 cell line. Asian Pac J Cancer Prev. 2012;13:3471–6.
43. Park JY, Kricka LJ, Fortina P. Next-generation sequencing in the clinic. Nat Biotechnol. 2013;31:990–2.
44. Xuan J, Yu Y, Qing T, Guo L, Shi L. Next-generation sequencing in the clinic: promises and challenges. Cancer Lett. 2013;340:284–95.
45. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. Nature. 2013;501:355–64.
46. Speicher MR. Single-cell analysis: toward the clinic. Genome Med. 2013;5:74.

## Publisher's Note