Identification of multidrug chemoresistant genes in head and neck squamous cell carcinoma cells

Neha Khera[a], Asvika Soodhalaagunta Rajkumar[a], Khlood Abdulkader M Alkurdi[a], Zhiao Liu[a,b], Hong Ma[b], Ahmad Waseem[a] and Muy-Teck Teh[a,b,1]

[a]Centre for Oral Immunobiology and Regenerative Medicine, Institute of Dentistry, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, England, United Kingdom.

[b]China-British Joint Molecular Head and Neck Cancer Research Laboratory, Affiliated Stomatological Hospital of Guizhou Medical University, Guizhou, China.

[1]Corresponding Author at:

Centre for Oral Immunobiology and Regenerative Medicine,
Barts and The London School of Medicine and Dentistry,
Queen Mary University of London,
The Blizard Building,
4, Newark Street,
London E1 2AT
United Kingdom.
Tel: +44 (0) 20 7882 7140
Fax: +44 (0) 20 7882 7137
Email: m.t.teh@qmul.ac.uk
ORCID iD: orcid.org/0000-0002-7725-8355

Keywords: multidrug chemoresistance, chemosensitization, cisplatin, drug-gene interaction, INHBA, NEK2, squamous cell carcinomas, oral cancer.

Abstract

Multidrug resistance renders treatment failure in a large proportion of head and neck squamous cell carcinoma (HNSCC) patients that require multimodal therapy involving chemotherapy in conjunction with surgery and/or radiotherapy. Molecular events conferring chemoresistance remain unclear. Through transcriptome datamining, 28 genes were subjected to pharmacological and siRNA rescue functional assays on 12 strains of chemoresistant cell lines each against cisplatin, 5-fluorouracil (5FU), paclitaxel (PTX) and docetaxel (DTX). Ten multidrug chemoresistance genes (TOP2A, DNMT1, INHBA, CXCL8, NEK2, FOXO6, VIM, FOXM1B, NR3C1 and BIRC5) were identified. Of these, four genes (TOP2A, DNMT1, INHBA and NEK2) were upregulated in an HNSCC patient cohort (n=221). Silencing NEK2 abrogated chemoresistance in all drug-resistant cell strains. INHBA and TOP2A were found to confer chemoresistance in majority of the drug-resistant cell strains whereas DNMT1 showed heterogeneous results. Pan-cancer Kaplan-Meier survival analysis on 21 human cancer types revealed significant prognostic values for INHBA and NEK2 in at least 16 cancer types. Drug library screens identified two compounds (Sirodesmin A and Carfilzomib) targeting both INHBA and NEK2 and re-sensitised cisplatin-resistant cells. We have provided the first evidence for NEK2 and INHBA in conferring chemoresistance in HNSCC cells and siRNA gene silencing of either gene abrogated multidrug chemoresistance. The two existing compounds could be repurposed to counteract cisplatin chemoresistance in HNSCC. This finding may lead to novel personalised biomarker-linked therapeutics that can prevent and/or abrogate chemoresistance in HNSCC and other tumour types with elevated NEK2 and INHBA expression. Further investigation is necessary to delineate their signalling mechanisms in tumour chemoresistance.

Background

Head and neck squamous cell carcinomas (HNSCCs) constitute 90% of all head and neck cancers. Whilst a minority of HNSCCs are caused by human papilloma virus (HPV) infection the majority of the more aggressive HPV-negative HNSCCs (75%) are associated with tobacco and alcohol use [1]. Although the cure rate of HNSCC patients with early-stage disease treated with primary surgery and/or radiotherapy has been excellent (70-90% 5-year overall survival) [1]; unfortunately, two-thirds of HNSCC patients present with advanced-stage disease suffer from poor survival outcome due to limited treatment options and/or treatment failure. Hence, the long-term survival rate of HNSCC patients remains unchanged over many decades at about 50% despite advancements in treatment modalities [1].

For HNSCC patients requiring multimodal therapy involving chemotherapy, cisplatin, 5-fluorouracil (5FU) and paclitaxel (PTX) or docetaxel (DTX) are amongst the most commonly used chemotherapeutic agents often used in combinations [1, 2]. Unfortunately, treatment failure due to development of resistance to chemo and/or radiotherapy remains a major cause of HNSCC poor survival rates. Unlike lung and breast cancer patients, all HNSCC patients are treated with almost the same combinations of treatment irrespective of the genetic makeup of their cancer. This is mainly due to poor understanding of molecular heterogeneity of HNSCC. Research into molecular biomarkers that can stratify sub-populations and indicate the most suitable intervention based on individual patient's tumour molecular profile would reduce toxicity, improve morbidity and treatment outcome [1, 3].

A number of key mechanisms for conferring intrinsic chemoresistance in HNSCC tumour cells have been studied and these include perturbations of pathways regulating apoptosis/cell death, DNA damage repair, epithelial mesenchymal transition, cell cycle, cancer stem cell, chromatin/epigenetic, miRNA processing, autophagy and stroma/matrix, immune cell interactions [4]. Although a number of molecular markers have been proposed for counteracting chemoresistance in HNSCC [4], exploitation of molecular markers for risk stratification in HNSCC patients prior to treatment decision largely remains at infancy [1, 3].

This study explored using a combination of bioinformatics transcriptome data mining, differential gene expression analysis in chemoresistant cell line models and validation in clinical HNSCC tumour specimens, pharmacological dose-response drug library screen and cell culture models with the aim to identify key multidrug-resistant biomarker genes and repurpose existing drugs to counteract chemoresistance.

Methods

All details of materials and methods can be found in Additional File 1. In brief, the following methods were used in this study: transcriptome data mining to identify differentially expressed genes, clinical HNSCC tissue cohort to validate candidate genes, cell culture models to validate gene expression and establish drug-resistant cell strains for functional analyses, cell viability assays to measure drug responses, pharmacological dose-response assays to identify drug-gene interactions, siRNA assays to validate candidate chemoresistant genes, reverse transcription quantitative PCR (RT-qPCR) to measure gene expression and drug library screens to identify potential existing known drugs to counteract HNSCC chemoresistance.

Results

Transcriptome data mining and gene selection

Meta-analyses of eight independent HNSCC microarray studies (see Additional File 2: Table S1) were performed using the cancer microarray database Oncomine (www.oncomine.org) to identify differentially expressed genes in studies comparing HNSCC with normal oral mucosa. Initially, top 40 differentially expressed genes were selected based on their reported P-values (>0.001). We performed RT-qPCR to quantify each of the 40 genes in a panel of eight primary normal human oral keratinocytes (OK355, HOKG, OK113, NOK, NOK1, NOK3, NOK16 and NOK376) and ten HNSCC cell lines (SCC4, SCC9, SCC15, SCC25, SqCC/Y1, UK1, VB6, CaLH2, CaDec12 and 5PT) to identify and validate differentially expressed genes. Of the 40 genes, 28 were found to be differentially expressed in our cell line panels and have been implicated in the regulation of matrix remodelling, immune modulation, cell proliferation & differentiation, stem cell renewal, epigenetic programming and genomic instability (Fig. 1A and Additional File 2: Table S2).

Identification of common multidrug-resistant genes

With an aim to identify key genes that mediate chemoresistance in molecularly different background, we have selected three cell lines to represent diverse molecular background from oral premalignancy (SVpgC2a), carcinogen (nicotine)-transformed malignancy (SVFN8) and a patient HNSCC tumour-derived malignancy (CALH2). In order to identify common multidrug-resistant genes, we generated four drug-resistant cell strains (R1: Cisplatin, R2: 5FU, R3: PTX and R4: DTX) for each of the three different cell lines (SVpgC2a, SVFN8 and CaLH2) giving rise to a total of 12 cell/drug-resistant combination strains (Fig. 1B). For each cell/drug-resistant combination strain, we challenged each wildtype (WT) and drug-resistant cell strain with the corresponding drug and measured the differential expression of the 28 genes by RT-qPCR to identify drug dose-dependent response genes to each chemotherapeutic drug (Additional File 2: Fig. S1-S12). We then performed statistical t-test and regression analyses on the 28 genes for each cell/drug-resistant strain to identify differentially expressed genes between drug-resistant and WT cells for each drug. To identify the most common drug-resistant genes, the 28 genes were ranked in descending order according to their frequency of occurrence as top significant genes across the whole panel of 12 cell/drug combinations. Of these, we selected the top four upregulated genes (TOP2A, DNMT1, INHBA and NEK2) across the entire cell/drug combinations to further investigate their roles in conferring multidrug resistance.

Differential gene expression in HNSCC clinical tissue samples

In order to confirm that these four genes (TOP2A, DNMT1, INHBA and NEK2) were indeed upregulated in HNSCC tumours, we performed RT-qPCR on a UK HNSCC tissue cohort to quantify their relative gene expression levels in adjacent margin (n=98) and HNSCC core tumour tissues (n=123). All four genes were confirmed to be significantly upregulated in HNSCC tumour compared to margin tissues ($P<10^{-5}$; Fig. 1C top panel). To cross validate our findings with an external cohort, we queried the four gene expression using the pan-cancer GEPIA 'Box Plot' tool based on transcriptomic data of The Cancer Genome Atlas (TCGA)/The Genotype-Tissue Expression (GTEx). In agreement, our findings are consistent with

TCGA/GTEx HNSCC cohort demonstrating significant upregulation of all four genes in HNSCC (n=519) over normal mucosa (n=44) samples (P<0.01; Fig. 1C bottom panel).

Reversal of chemoresistance by siRNA gene silencing

To investigate if the four genes (TOP2A, DNMT1, INHBA and NEK2) identified above were conferring multidrug resistance across the 12 different cell strains, we performed gene silencing using siRNA to knockdown each of these genes in both the WT and drug-resistant cells in response to each corresponding drug. We hypothesised if the genes were necessary to sustain drug resistance, knockdown of the genes would abrogate chemoresistance. To test this hypothesis, we transfected gene-specific siRNA and treated both WT and drug-resistant cells to serial-dilutions of corresponding drug to determine their IC_{50} values (drug potency). Abrogation of chemoresistance would result in a shift in IC_{50} values of resistant cells towards IC_{50} of WT cells (i.e., reducing the fold difference between the two IC_{50} values). We included untransfected (+H_2O, containing transfection reagent only) and control siRNA (siCTRL) as controls and confirmed gene-specific siRNA silencing by RT-qPCR (Additional File 2: Fig. S13). We screened for reversal of chemoresistance by siRNA against each of the four genes in the three cell lines (SVpgC2a, SVFN8 and CaLH2) each with drug-resistance to each of the four chemotherapeutic drugs (cisplatin, 5FU, PTX and DTX). The chemosensitivity (IC_{50}-fold change between resistant and WT cells) of all the 12 cell strains are summarised in Fig. 1D and individual dose-response curves data are shown in Additional File 2: Fig. S14-S16. Gene expression levels and IC_{50} values in siCTRL transfected cells were very similar to untransfected cells indicating that siCTRL did not induce any non-specific or off-target effects. In SVpgC2a cells, siTOP2A and siNEK2 both completely reversed chemoresistance (P<0.001) in all four drug-resistant strains. siDNMT1 reversed only PTX- and DTX-resistant strains whilst, siINHBA reversed only cisplatin and PTX-resistant strains. In SVFN8 cells, similar to results from SVpgC2a cells, siTOP2A and siNEK2 completely reversed drug resistance in all four drug-resistant strains. However, unlike results for SVpgC2a, siDNMT1 and siINHBA showed partial reversal of resistance in all four different drug-resistant strains. Interestingly, in CaLH2 cells (HNSCC tumour-derived cell line), siNEK2 and siINHBA showed complete reversal of drug resistance in all four drug-resistant strains, whilst siDNMT1 and siTOP2A showed only partial reversal of chemoresistance.

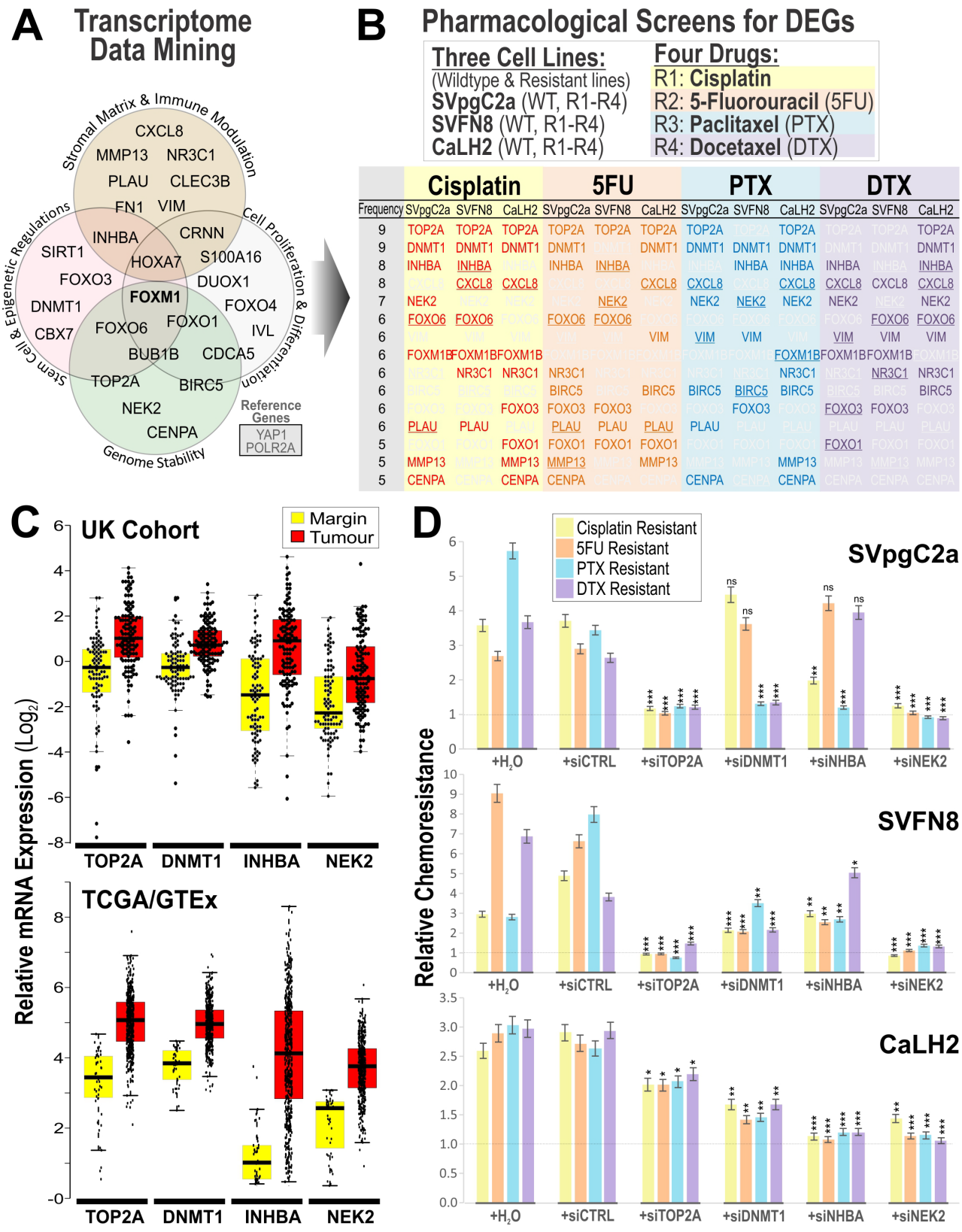


Fig. 1. *Identification and validation of candidate genes responsible for conferring multidrug resistance in HNSCC.* ***A****, Transcriptome data mining from eight independent gene expression microarray studies (comparing HNSCC tumour and normal oral tissues samples) identified 28 short-listed genes involved in the regulation of matrix remodelling, immune modulation, cell proliferation & differentiation, stem cell renewal, epigenetic programming and genomic instability.* ***B****, Pharmacological dose-response screening for multidrug-resistant differentially expressed genes in three cell lines (SVpgC2a, SVFN8 and CaLH2) each between parental WT cells and their corresponding drug-resistant strains*

for four chemotherapeutic drugs (R1: Cisplatin, R2: 5FU, R3: PTX and R4: DTX). For each cell/drug resistant combination strain, differential expression of the 28 genes were measured by RT-qPCR to identify drug dose-dependent response genes (see Additional File 2: Fig. S1-S12). Table shows the top differentially expressed genes in corresponding drug-resistant cell strains are shown in coloured (statistically significant) gene symbols (non-significant genes in white text). Genes with underlines indicate downregulation in drug-resistant cells, otherwise, upregulation. The list of 28 genes were ranked in descending order according to their frequency of occurrence as top significant genes across the whole panel of 12 cell/drug combinations. ***C****, Validation of gene expression levels of TOP2A, DNMT1, INHBA and NEK2 in two different HNSCC patient cohorts: Top panel, a UK cohort with adjacent margin (n=98) and HNSCC tumour core tissues (n=123). The relative mRNA expression levels of each of the four genes were measured using RT-qPCR against two reference genes (YAP1 and POLR2A) measured in duplicate wells. Data were plotted as beeswarm dot-plot with box-and-whisker overlays (minimum, box: median, and 25-75%, percentiles and maximum). Statistical t-test were performed between the margin and tumour samples and all four genes showed* $P<1x10^{-5}$*. Bottom panel: Differential expression of the four genes in HNSC cohort from TCGA/GTEx transcriptomic data comparing margin (n=44) and HNSCC (n=519) were all significantly upregulated in tumour (P<0.01, one-way ANOVA).* ***D****, Effects of siRNA gene silencing of TOP2A, DNMT1, INHBA and NEK2 on reversal of chemoresistance (or re-sensitisation). Summary of relative chemoresistance following gene-specific siRNA knockdown in SVpgC2a, SVFN8 and CaLH2 cells, each resistant to either cisplatin, 5FU, PTX or DTX. Relative chemoresistance was calculated as fold-change between* IC_{50} *of drug-resistant cells and* IC_{50} *of corresponding WT cells.* IC_{50} *drug potency values of each of the four chemotherapeutic drugs on WT and drug-resistant cells were measured using crystal violet cell viability assay (Additional File 2: Fig. S14-S16). Statistical t-test was performed between controls (mock transfection/+*H_2O *and siCTRL were combined as one group) vs each of the gene-specific siRNA and their corresponding P-values are indicated (*<0.05; **<0.01; ***<0.001; ns, not significant) within the charts.*

Pan-cancer Kaplan-Meier prognostic biomarker meta-analysis

To further investigate the four genes (TOP2A, DNMT1, INHBA and NEK2) differential expression and if they have any prognostic value in different human cancer types, we performed data mining on publicly available pan-cancer databases Oncomine and KM Plotter with RNA-seq transcriptome (KM-Plotter.com) containing 54,675 genes with survival outcome for 21 different human cancer types (Additional File 3). All four genes were found to be upregulated in the majority of human cancer types with few minor exceptions (Fig. 2A). Upregulation of INHBA, NEK2, TOP2A and DNMT1 were associated with poor prognosis in 13 (87%), 9 (82%), 7 (64%) and 3 (38%) of tumour types, respectively (Fig. 2B-2C). For HNSCC, upregulation of INHBA, NEK2 and TOP2A but not DNMT1 were associated with poor prognosis (Fig. 2D). All four marker (alone and in combinations) were examined for their synergistic prognostic values for each cancer type and the most significant prognostic marker (or markers in combinations) are shown in Fig. 2E. Overall, INHBA followed by NEK2 appeared to be pan-cancer prognostic markers for predicting poor survival outcome in the majority of cancer types. INHBA (alone and/or in combinations with TOP2A, NEK2 or DNMT1) predicted poor prognosis in 16 out of 21 different human cancer types, including HNSCCs (Fig. 2E). These data are consistent with data found in another pan-cancer database Gene Expression Profiling Interactive Analysis (GEPIA) based on the Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) RNA-seq data for at least 33 human cancer types (Additional File 1).

Identification and repurposing drugs targeting INHBA and NEK2 for counteracting cisplatin resistance

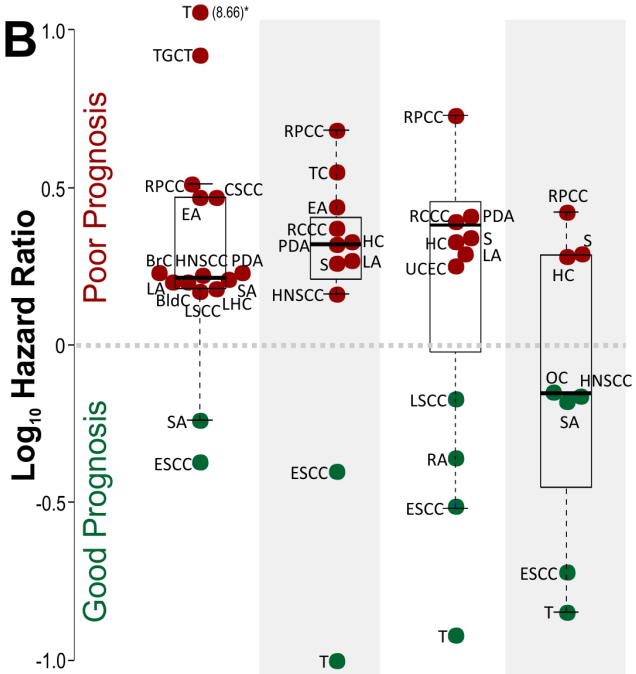
We performed a cell viability drug screen on a total of 537 compounds consisting of 147 approved oncology drugs (AOD IX) and 390 natural products (Set V) in WT and cisplatin-resistant (CR) CaLH2 with an aim to identify and repurpose existing drugs that suppress INHBA and/or NEK2 gene expression to counteract chemoresistance in HNSCC. We chose the CaLH2 cell line to investigate the effect of candidate drugs as both INHBA and NEK2 genes were not found to be differentially expressed between WT and CR cells in respond to cisplatin (Fig. 1B). The initial viability screening results led us to select nine most effective compounds (P<0.05), of which three compounds (D1-D3) were selected as control drugs with specificity for killing WT but not CR cells. The next three compounds (D4-D6) were specific for killing CR cells and the remaining three compounds (D7-D9) killed both WT and CR cells (Additional File 2: Fig. 17A). To investigate if the nine compounds were capable of inhibiting INHBA and NEK2 gene expression in a dose-dependent manner in respective WT and CR cells, we treated cells with serial dilution of each of the nine compounds and measured relative gene expression levels of INHBA and NEK2 using RT-qPCR. The first three control

compounds (D1-D3), consistent with cell viability results, showed dose-dependent inhibition on both INHBA and NEK2 expression in WT cells but not in CR cells. Of the remaining drugs, D4 and D7 dose-dependently inhibited with comparable potency on both INHBA (Fig. 2F) and NEK2 (Fig. 2G) gene expression in both WT and CR cells. D5 inhibited only INHBA but not NEK2. D9 did not show dose-dependent gene inhibition in either WT or CR cells (Additional File 2: Fig. 17B-17C). Potency (IC_{50}) of D4, D5 and D7 on gene inhibition were found to range from $1.1x10^{-8}$ M to $6.3x10^{-7}$. As D5 only inhibited INHBA and not NEK2, this compound was not further investigated. The chemical structures and identities of the two selected compounds D4 (Sirodesmin A) and D7 (Carfilzomib) are shown in Fig. 2H (ID of other compounds are shown in additional File 2: Fig. 17D). Subsequent cisplatin dose-dependent cell viability assays in the presence of a single dose (1 μM) of either D4 or D7 re-sensitised (leftward shift in dose-respond curves) the potency (IC_{50}) of cisplatin from 15.10 ± 1.59 μM to 0.52 ± 0.12 μM (29-fold by D4; t-test $P=1.7\times10^{-5}$) or to 0.54 ± 0.07 μM (27.7-fold by D7; $P=3.6\times10^{-6}$), respectively. This demonstrated that cisplatin-resistant cells could be significantly re-sensitised to cisplatin by addition of either D4 or D7 (Fig. 2I). We noted a biphasic cell viability cisplatin dose-response curves in the presence of either D4 or D7 (1 μM); this may indicate the involvement of multiple mechanisms of cisplatin resistance. Both D4 or D7 showed partial dose-dependent sensitisation only within the lower doses (0.1-1 μM) of cisplatin but plateaued in higher doses (1-10 μM) of cisplatin. We speculated that this could be due to the presence of different populations of cells (e.g., EMT cells and stem cells) within the culture and/or that there are multiple distinct signalling pathways involved. D4 and D7 could be acting only on one of these cisplatin-resistant pathways, hence demonstrating a biphasic response. Further investigation is required to delineate these mechanisms.

A

Analysis Type by Cancer	INHBA Cancer vs. Normal		NEK2 Cancer vs. Normal		TOP2A Cancer vs. Normal		DNMT1 Cancer vs. Normal	
Bladder Cancer	1		1		6		2	
Brain and CNS Cancer			2		10		1	
Breast Cancer	11		14	1	18	1		1
Cervical Cancer	1		4		4		3	
Colorectal Cancer	16		12		18		3	
Esophageal Cancer	2		2		2		3	
Gastric Cancer	7		3		6		1	
Head and Neck Cancer	5		1		8			
Kidney Cancer		1		5	3			
Leukemia		8	1	2		5	1	1
Liver Cancer			4		4		2	
Lung Cancer			11		17		1	
Lymphoma			3		7			
Melanoma		2			1			
Myeloma								
Other Cancer	1		3	4	4	1	2	1
Ovarian Cancer	2		2		4			
Pancreatic Cancer	2		1		2			
Prostate Cancer				1	1			
Sarcoma	2		5		11		5	
Significant Unique Analyses	50	11	68	13	125	7	24	3
Total Unique Analyses	439		445		462		454	

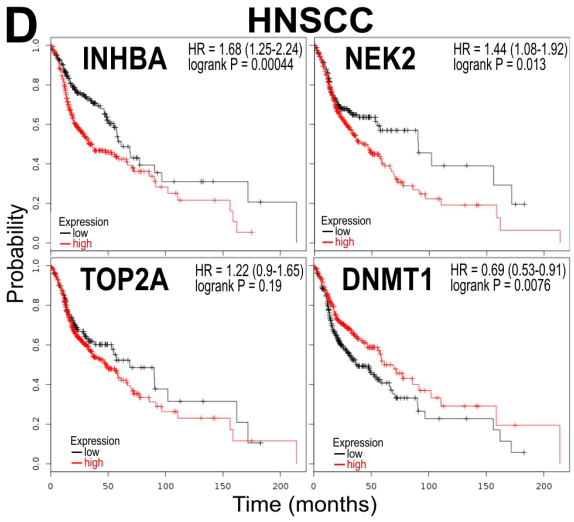
B



C

	INHBA	NEK2	TOP2A	DNMT1
Poor Prognosis	13 (87%)	9 (82%)	7 (64%)	3 (38%)
Good Prognosis	2 (13%)	2 (18%)	4 (36%)	5 (63%)
Total # Cancer	15 (100%)	11 (100%)	12 (100%)	8 (100%)

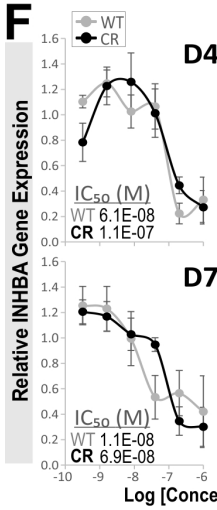
D



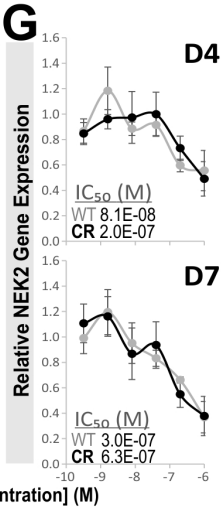
E Most Significant Prognostic Marker(s)

Marker(s)	HR	logrank P	Abbr.	Cancer Type
INHBA	1.58	4E-03	BldC	Bladder carcinoma (n=405)
INHBA	1.70	2E-03	BrC	Breast cancer (n=1090)
INHBA	2.94	2E-05	CSCC	Cervical squamous cell carcinoma (n=304)
I+N+D	2.65	6E-03	EA	Esophageal adenocarcinoma (n=80)
I+T+D	0.27	1E-03	ESCC	Esophageal squamous cell carcinoma (n-81)
I+N	1.65	2E-04	HNSCC	Head & neck squamous cell carcinoma (n=500)
N+T	2.53	7E-10	RCCC	Renal clear cell carcinoma (n=530)
I+N+T+D	6.21	2E-11	RPCC	Renal papillary cell carcinoma (n=288)
I+N+T	2.29	2E-06	HC	Hepatocellular carcinoma (n=371)
N+T	1.94	4E-05	LA	Lung adenocarcinoma (n=513)
I+N	1.53	2E-03	LSCC	Lung squamous cell carcinoma (n=501)
DNMT1	0.71	2E-02	OC	Ovarian cancer (n=374)
T+D	2.61	9E-06	PDA	Pancreatic ductal adenocarcinoma (n=117)
T+D	8.69	2E-02	PP	Pheochromocytoma & Paraganglioma (n=178)
I+T+D	0.33	4E-03	RA	Rectum adenocarcinoma (n=165)
I+T+D	2.20	2E-04	S	Sarcoma (n=259)
INHBA	1.63	4E-03	SA	Stomach adenocarcinoma (n=375)
INHBA	8.30	3E-02	TGCT	Testicular germ cell tumour (n=134)
INHBA	5E+08	2E-04	T	Thymoma (n=119)
I+N+T	3.74	7E-03	TC	Thyroid carcinoma (n=502)
I+T+D	1.56	4E-03	UCEC	Uterine corpus endometrial carcinoma (n=543)

F

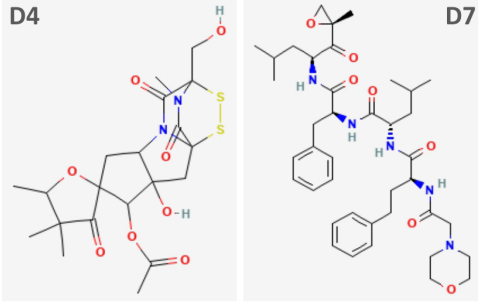


G



H

Sirodesmin A		Carfilzomib
294408	NSC	758252
52988-50-8	CAS	868540-17-4
146286	PubChem SID	440803622
325517	PubChem CID	11556711
$C_{20}H_{26}N_2O_8S_2$	Mol. Formula	$C_{40}H_{57}N_5O_7$
486.6	Mol. Weight	719.9



I

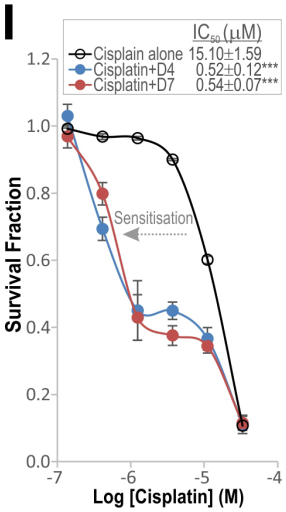


Fig. 2. *Pan-cancer bioinformatics data mining and prognostic analysis.* ***A****, Bioinformatics data mining from Oncomine databases on differential gene expressions of INHBA, NEK2, TOP2A and DNMT1 across 20 different human cancer types as indicated. The number within each coloured box indicates the number of significant unique studies. Red and blue colours indicate gene expression upregulation and downregulation, respectively. Cell colour scale is determined by the best gene rank percentile for the analyses (dark red/blue = top 1%; red/blue = top 5%; pale red/blue = top 10%).* ***B****, Kaplan-Meier RNA-seq transcriptome prognostic analysis for INHBA, NEK2, TOP2A and DNMT1 on 21 different human cancer types. Hazard ratios (with logrank P<0.05) extracted from KM-plotter database were plotted here as beeswarm dot-plot with box-and-whisker overlays (minimum, box: median, and 25-75%, percentiles and maximum) to demonstrate individual marker prognostic value for each cancer type. Dark red indicates marker associated with poor prognosis and green for markers associated with good prognosis (abbreviations listed in panel E). *Note: outlier (Thymoma, Log_{10} HR = 8.66) was plotted outside the chart for reference.* ***C****, Table listing corresponding number (and %) of cancer type analysed for each marker with poor or good prognosis.* ***D****, Individual Kaplan-Meier plots for INHBA, NEK2, TOP2A and DNMT1 in HNSCC tumour samples (n=500) with hazard ratio (HR) and logrank P values as shown within each panel.* ***E****, All four markers (alone and in combinations) were examined for their synergistic prognostic values for each cancer type, the most significant prognostic marker (or markers in combinations) are tabulated here. Hazard ratio (HR) values were shown with colour scales applied to indicate poor (dark red) or good (green) prognosis with their corresponding logrank P values (colour scales indicate their relative levels of significance) for each cancer type (n = the number of samples in each cancer type). Marker abbreviations: I, INHBA; N, NEK2; T, TOP2A and D, DNMT1).* ***F-I****, Drug library screen to identify drug-gene interactions for counteracting chemoresistance in HNSCC cells. Nine compounds (D1-D9; Additional File 2: Fig. 17) were selected. Shown here are two compounds (D4, and D7) with dose-dependent inhibition on both INHBA (****F****) and NEK2 (****G****) gene expression in WT and CR CaLH2 cells. Each datapoint represents relative gene expression (mean ± SEM) of quadruplicates quantified using RT-qPCR. Drug potencies (IC_{50}) on respective gene inhibition are displayed within each panel.* ***H****, Chemical structure and identity of compounds D4 and D7.* ***I****, Re-sensitisation of CR CalH2 cells by addition of D4 or D7 (1 μM) to cisplatin dose-response measured using AlamarBlue cell viability assay. Each datapoint represents a mean ± SEM of n=6 replicates. Cisplatin potency values (mean IC_{50} ± SEM of n=6) in the absence or presence of either D4 or D7 single concentration (1 μM) are shown within the figure. Statistical t-test was performed between cisplatin alone vs cisplatin+D4 or cisplatin+D7 and their corresponding P-values are indicated within the figure as ***<0.001.*

Discussion

Currently there are no specific molecular biomarkers that can indicate which HNSCC patient is susceptible to developing drug resistance. This study presented a series of experiments using a combination of bioinformatics data mining, cell culture model and pharmacological approaches to identify key transcriptome biomarkers that confer multidrug chemoresistance. We have identified ten candidate genes (TOP2A, DNMT1, INHBA, CXCL8, NEK2, FOXO6, VIM, FOXM1B, NR3C1 and BIRC5) of which, four (TOP2A, DNMT1, INHBA and NEK2) were confirmed to be significantly upregulated in our UK HNSCC tumour cohort and also consistent with TCGA/GTEx data. NEK2, DNMT1 and FOXM1B had been previously shown to be upregulated in HNSCC tumour cohorts from UK, Norway and China [5, 6]. Furthermore, TOP2A, DNMT1, INHBA, CXCL8, NEK2, NR3C1 and BIRC5 were previously shown to be differentially expressed in independent HNSCC patient cohorts from UK, China and India [7], and these genes were part of a multigene biomarker panel for molecular diagnosis of HNSCC and risk stratification in dysplastic oral premalignant disorders [7]. Nevertheless, their roles in HNSCC chemoresistance remain unclear.

Here, we uncovered an essential role for a cell cycle gene, NEK2 (Never in mitosis gene A-related kinase 2), in conferring multidrug chemoresistance in HNSCC cells whereby targeted siRNA gene silencing against NEK2 led to complete abrogation of chemoresistance in all 12 chemoresistant cell strains (to four different drugs: cisplatin, 5FU, PTX and DTX). Consistently, our pan-cancer Kaplan-Meier survival analysis showed that upregulation of NEK2 predicted poor prognosis in HNSCC patients. Although NEK2 has been previously reported to confer chemoresistance in multiple human malignancies, to our knowledge we presented the first evidence for NEK2 in conferring multidrug chemoresistance in HNSCC. We have also presented evidence for INHBA (inhibin subunit beta A), TOP2A (DNA topoisomerase II alpha) and DNMT1 (DNA methyltransferase 1) in conferring chemoresistance in the majority of drug-resistant cell strains. Targeted siRNA on each of these genes showed mixed responses across the 12 chemoresistance cell strains perhaps due to inherent heterogeneity of the different parental cell lines. Further investigation is required to delineate and differentiate their molecular pathways in these cell strains.

In support of our HNSCC cohort data, previous studies have also demonstrated differential upregulation of INHBA in HNSCC [8]. Our pan-cancer Kaplan-Meier survival analysis further revealed that INHBA (alone and/or in combinations with TOP2A, NEK2 or DNMT1) predicted poor prognosis in 16 out of 21 different human cancer types, including HNSCC. In agreement with a role in chemoresistance found in this study, INHBA has been shown to be part of a 7-gene prognostic signature that predicts the outcome of HNSCC patients treated with postoperative radio(chemo)therapy [9]. Given that INHBA gene encodes a member of the TGF-beta (transforming growth factor-beta) superfamily of proteins which involve in the regulations of EGFR (epidermal growth factor receptor) [10] and oncogenic transcription factor RUNX2 [11] pathways in HNSCC cells, highlights the significance of INHBA as an important novel molecular target and prognostic biomarker for HNSCC.

With an aim to repurpose the use of licensed drugs for counteracting cisplatin resistance in HNSCC, from our drug-gene interaction library screens, we identified two drugs (Sirodesmin A and Carfilzomib) targeted both INHBA and NEK2 in a dose-dependent manner and re-sensitised cisplatin resistant cells. Sirodesmin A is a natural metabolite produced by the fungus Sirodesmium diversum (ascomycete fungi) [12] and little is known about its activity on human cancer cells. To our knowledge, we presented the first evidence that Sirodesmin A counteracted cisplatin resistance in a HNSCC cell line and dose-dependently inhibited both INHBA and NEK2 gene expression. Carfilozomib (Kyprolis®), a derivative of a bacterial actinomycete irreversible proteosome inhibitor epoxomicin, is licensed for treating patients with relapsed and/or refractory multiple myeloma [13]. Whilst Carfilozomib has been shown to potentiate the effect of cisplatin in a number of cancer types such as multiple myeloma [14], ovarian [15] and neuroblastoma [16], our results demonstrated for the first time that Carfilozomib dose-dependently inhibited both INHBA and NEK2 gene expression, and re-sensitise cisplatin resistant HNSCC cells. We hypothesised that Sirodesmin A or Carfilozomib could be repurposed to counteract cisplatin resistance in tumours with elevated NEK2 and/or INHBA gene expression. Further investigations are necessary to delineate their drug-gene interactions and mechanism of actions in counteracting chemoresistance.

Conclusions

We presented the first evidence for NEK2 in conferring multidrug chemoresistance in HNSCC cells and that targeted siRNA gene silencing led to complete reversal of drug resistance to cisplatin, 5FU, PTX and DTX. INHBA and TOP2A were found to confer chemoresistance in the majority of drug-resistant cell strains whereas DNMT1 showed heterogeneous effects on chemoresistance. Pan-cancer Kaplan-Meier survival analysis on 21 human cancer types revealed significant prognostic values for NEK2 and INHBA in the majority of cancer types. We further identified a naturally occurring fungal derivative Sirodesmin A and a licensed anticancer drug Carfilzomib, both targeting NEK2 and INHBA, could re-sensitise resistant HNSCC cells to cisplatin. This finding requires further investigations into the potential of repurposing licensed drugs for reversing chemoresistance in HNSCC patients.

Abbreviations

5FU	5-fluorouracil
cDNA	complementary DNA
CR	cisplatin resistant cells
DMEM	Dulbecco's Modified Eagle Medium
DNMT1	DNA methyltransferase 1
DTX	docetaxel
FOXM1B	forkhead box M1 isoform B
IC_{50}	Drug potency, concentration of drug which induced 50% inhibition
HNOK	human normal oral keratinocytes
HNSCC	head and neck squamous cell carcinoma
HPV	human papilloma virus

INHBA	inhibin subunit beta A
KM	Kaplan-Meier
miRNA	micro-RNA
NEK2	Never in mitosis gene A-related kinase 2
NHS	National Health Service
PD-1	programmed cell death 1
PTX	paclitaxel
QMUL	Queen Mary University of London
RT	room temperature
RT-qPCR	reverse transcription quantitative polymerase chain reaction
SCC	squamous cell carcinoma
siRNA	short interfering RNA
SV40T	Simian virus 40T
SVFN8	a transformed/malignant cell line derived from SVpgC2a
SVpgC2a	premalignant SV40T-antigen immortalised human buccal keratinocytes
TGF-beta	transforming growth factor-beta
TOP2A	DNA topoisomerase II alpha
WT	wildtype cells

Declarations

Ethics approval and consent to participate – “The use of fresh clinical specimens collected in the UK was approved by the NHS Research Ethics Committee (06/MRE03/69). All tissue samples were previously collected according to local ethical committee-approved protocols and informed patient consent was obtained from all participants”.

Consent for publication – Not applicable

Availability of data and material – “All data generated or analysed during this study are included in this published article [and its supplementary information files]”.

Competing interests - "The authors declare that they have no competing interests”

Funding - Affiliated Stomatological Hospital of Guizhou Medical University

Authors' contributions - Author contributions: M.T.T. designed research; N.K, A.S.R, K.A.M.A, Z.L., M.T.T. performed research; A.W., H.M. contributed new reagents/analytic tools; N.K, A.S.R, K.A.M.A, Z.L., M.T.T. analysed data; and M.T.T. wrote the paper.

Acknowledgements - This work was supported by Affiliated Stomatological Hospital of Guizhou Medical University (to Z.L., H.M. and M.T.T.). The authors are thankful to the Centre for Immunobiology and Regenerative Medicine (COIRM) for supporting this study.

References

1. Cramer JD, Burtness B, Le QT, Ferris RL: **The changing therapeutic landscape of head and neck cancer.** *Nat Rev Clin Oncol* 2019.
2. Gau M, Karabajakian A, Reverdy T, Neidhardt EM, Fayette J: **Induction chemotherapy in head and neck cancers: Results and controversies.** *Oral Oncol* 2019, **95:**164-169.
3. Alsahafi E, Begg K, Amelio I, Raulf N, Lucarelli P, Sauter T, Tavassoli M: **Clinical update on head and neck cancer: molecular biology and ongoing challenges.** *Cell Death Dis* 2019, **10:**540.
4. Usman S, Waseem NH, Nguyen TKN, Mohsin S, Jamal A, Teh MT, Waseem A: **Vimentin Is at the Heart of Epithelial Mesenchymal Transition (EMT) Mediated Metastasis.** *Cancers (Basel)* 2021, **13**.
5. Teh MT, Hutchison IL, Costea DE, Neppelberg E, Liavaag PG, Purdie K, Harwood C, Wan H, Odell EW, Hackshaw A, Waseem A: **Exploiting FOXM1-orchestrated molecular network for early squamous cell carcinoma diagnosis and prognosis.** *Int J Cancer* 2013, **132:**2095-2106.
6. Ma H, Dai H, Duan X, Tang Z, Liu R, Sun K, Zhou K, Chen H, Xiang H, Wang J, et al: **Independent evaluation of a FOXM1-based quantitative malignancy diagnostic system (qMIDS) on head and neck squamous cell carcinomas.** *Oncotarget* 2016, **7:**54555-54563.
7. Teh MT, Ma H, Liang YY, Solomon MC, Chaurasia A, Patil R, Tekade SA, Mishra D, Qadir F, Yeung JS, et al: **Molecular Signatures of Tumour and Its Microenvironment for Precise Quantitative Diagnosis of Oral Squamous Cell Carcinoma: An International Multi-Cohort Diagnostic Validation Study.** *Cancers (Basel)* 2022, **14**.
8. Shimizu S, Seki N, Sugimoto T, Horiguchi S, Tanzawa H, Hanazawa T, Okamoto Y: **Identification of molecular targets in head and neck squamous cell carcinomas based on genome-wide gene expression profiling.** *Oncol Rep* 2007, **18:**1489-1497.
9. Schmidt S, Linge A, Zwanenburg A, Leger S, Lohaus F, Krenn C, Appold S, Gudziol V, Nowak A, von Neubeck C, et al: **Development and Validation of a Gene Signature for Patients with Head and Neck Carcinomas Treated by Postoperative Radio(chemo)therapy.** *Clin Cancer Res* 2018, **24:**1364-1374.
10. Tsai CN, Tsai CL, Yi JS, Kao HK, Huang Y, Wang CI, Lee YS, Chang KP: **Activin A regulates the epidermal growth factor receptor promoter by activating the PI3K/SP1 pathway in oral squamous cell carcinoma cells.** *Sci Rep* 2019, **9:**5197.
11. Chang WM, Lin YF, Su CY, Peng HY, Chang YC, Lai TC, Wu GH, Hsu YM, Chi LH, Hsiao JR, et al: **Dysregulation of RUNX2/Activin-A Axis upon miR-376c Downregulation Promotes Lymph Node Metastasis in Head and Neck Squamous Cell Carcinoma.** *Cancer Res* 2016, **76:**7140-7150.
12. Curtis PJ, Greatbanks D, Hesp B, Cameron AF, Freer AA: **Sirodesmins A, B, C, and B, antiviral epipolythiopiperazine-2,5-diones of fungal origin: X-ray analysis of sirodesmin A diacetate.** *J Chem Soc Perkin 1* 1977, **2:**180-189.
13. Siegel DS: **From clinical trials to clinical practice: single-agent carfilzomib adverse events and their management in patients with relapsed and/or refractory multiple myeloma.** *Ther Adv Hematol* 2013, **4:**354-365.
14. Chao A, Wang TH: **Molecular mechanisms for synergistic effect of proteasome inhibitors with platinum-based therapy in solid tumors.** *Taiwan J Obstet Gynecol* 2016, **55:**3-8.
15. Zarei S, Reza JZ, Jaliani HZ, Hajizadeh MR, Sargazi S, Hosseinian H: **Effects of carfilzomib alone and in combination with cisplatin on the cell death in cisplatin-sensitive and cisplatin-resistant ovarian carcinoma cell lines.** *Bratisl Lek Listy* 2019, **120:**468-475.
16. Lee SI, Jeong YJ, Yu AR, Kwak HJ, Cha JY, Kang I, Yeo EJ: **Carfilzomib enhances cisplatin-induced apoptosis in SK-N-BE(2)-M17 human neuroblastoma cells.** *Sci Rep* 2019, **9:**5039.

Materials and Methods

Clinical Samples

The use of fresh clinical specimens collected in the UK was approved by the NHS Research Ethics Committee (06/MRE03/69). All tissue samples were previously collected according to local ethical committee-approved protocols and informed patient consent was obtained from all participants [1-3]. Fresh tissue biopsies were preserved in RNALater (#AM7022, Ambion, Applied Biosystems, Warrington, UK) and stored short-term at 4°C (1-7 days) prior to transportation and subsequent storage at -20°C until used. All frozen samples were digested with nuclease-free proteinase K at 60°C prior to mRNA extraction (Dynabeads mRNA Direct kit, Invitrogen).

Cell Culture & Establishment of Drug-resistant Cells

All cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM) with 10% foetal bovine serum and 1% penicillin/streptomycin and maintained in a humidified incubator with 5% CO_2/95% atmospheric air at 37°C. All primary normal human oral keratinocytes (OK355, HOKG, OK113, NOK, NOK1, NOK3, NOK16 and NOK376) were extracted from normal oral mucosa donated by healthy disease-free individuals undergoing wisdom tooth extraction and cultured as previously described [4, 5]. SVpgC2a was a non-transforming/premalignant Simian virus 40T (SV40T)-antigen immortalised human buccal keratinocytes [6] and SVFN8 was a carcinogen (nicotine) transformed/malignant cell line derived from SVpgC2a [4]. HNSCC cell lines (SCC4[7], SCC9[7], SCC15[7], SCC25[7], SqCC/Y1[8], UK1[9], VB6[9], CaLH2[9], CaDec12[9] and 5PT[9]) are all well characterised lines and were cultured as described previously [4, 5, 10]. The p53 mutational status of the three cell lines (SVpgC2a, SVFN8 and CaLH2) used in the current study for generating chemoresistant lines is not known. However, according to our previous transcriptome profiling data (NCBI's GEO database GSE89217 [11]), all these three cell lines showed undetectable level of p53 gene expression. Several mechanisms including gene silencing, gene deletion, mutations that lead to premature stop codon, etc., could contribute to undetectable p53 mRNA levels. As these are beyond the scope of the current study, future sequencing experiments would need to be conducted to identify the type of mutations involved in these cell lines.

Crystal Violet Cell Viability Assay

Crystal violet cell viability assay was performed in 96-well plates (Figure S1). Growth medium was aspirated and 30 µL/well of crystal violet solution (0.5% crystal violet in 30% ethanol) was added and incubated for 10 minutes at room temperature (RT). Cells were then washed with 200 µL/well distilled water prior to the addition of 100 µL/well of 1% SDS and incubated for 30 minutes at RT prior to absorbance measurement at 595 nm using a CLARIOstar microplate reader.

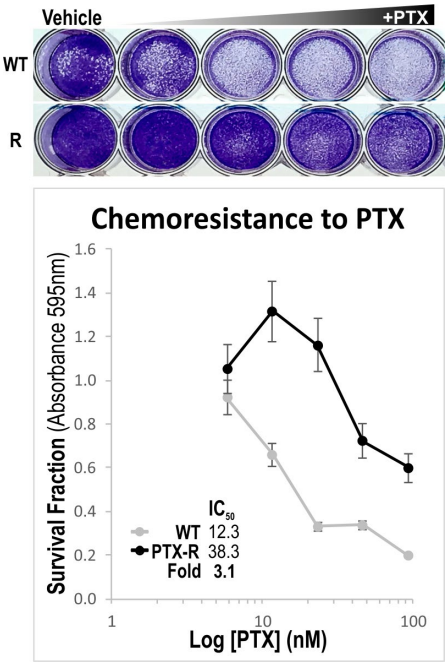


Figure S1. *Chemosensitivity assays for wildtype (WT) and drug-resistant cells measured by crystal violet cell viability assay following 72 hr drug incubation. An example shown here using WT and PTX-resistant (PTX-R) CaLH2 cell lines. IC_{50} values (shown here in nM) representing the degree of chemosensitivity for WT and PTX-R cells were determined using sigmoid-curve fitting algorithm on respective data points plotted as logarithmic drug concentrations on the X-axis and survival fraction (Absorbance at 595 nm) on the Y axis. IC_{50}-fold difference was calculated between WT and PTX-R cells as indicated within the graph.*

Establishment of Drug-resistant Cell Strains

Crystal violet cell viability dose-response curve (kill-curve) assays were first performed to determine the half maximal growth inhibition (IC_{50}) concentrations of cisplatin, 5-fluorouracil (5FU), paclitaxel (PTX) and docetaxel (DTX) on SVpgC2a, SVFN8 and CaLH2 cell lines. Each cell type was then cultured in growth media containing IC_{50} concentrations of each drug with regular changes of growth media containing freshly diluted drugs. When cells were proliferating, a 3-fold higher concentration above the IC_{50} of each drug were then added to the growth medium. This process was repeated until cells were able to proliferate in the highest drug concentrations over a period of ~6 months. Drug-resistant cells were then expanded in drug-free growth medium to create aliquots for cryopreservation until used for experiments. We have established a total of 12 drug resistant strains: four drug-resistant (Cisplatin, 5FU, PTX and DTX) strains for each of the three cell types (SVpgC2a, SVFN8 and CaLH2), with a minimum of 3-fold higher IC_{50} values than their corresponding parental wildtype cells Additional File 2: Figure S14-S16).

Drug-dependent Chemoresistant Gene Expression Assay

To identify genes that are differentially expressed when challenged with increasing doses of chemotherapeutic drugs in WT and its corresponding drug-resistant strains, cells were seeded into 96-well plate (8,000 cells/well) one day prior to drug treatment. Cells were then treated with serial dilutions of each drug for 24h prior to harvest for RT-qPCR to quantify the relative mRNA expression levels of 28 target genes (Figure S2).

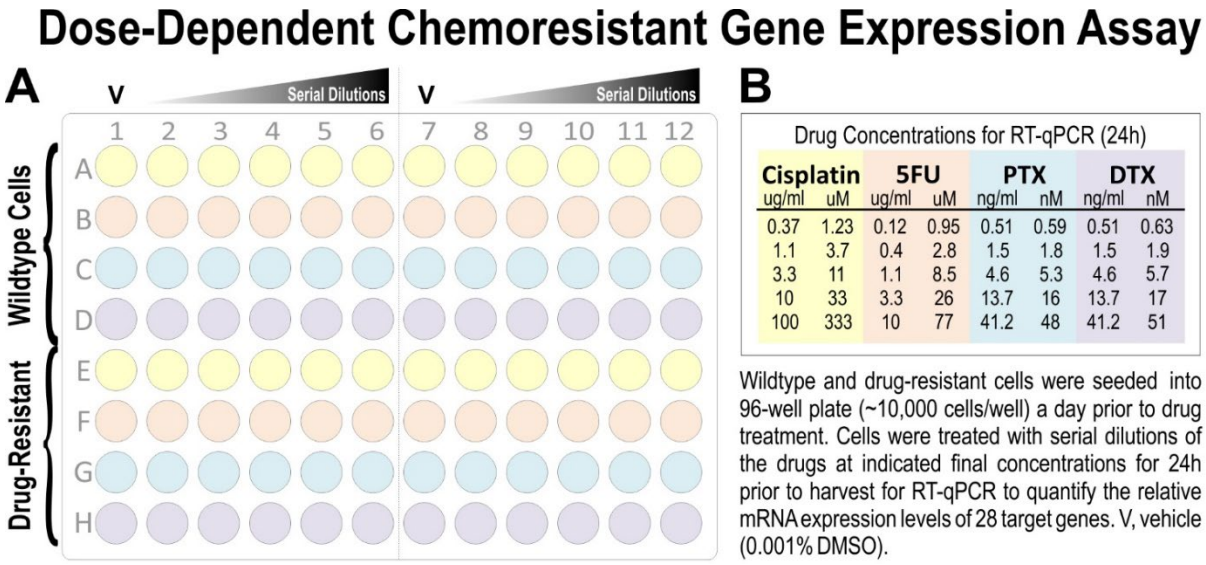


Figure S2. *Dose-dependent chemoresistant gene expression assay on wildtype and drug-resistant cells treated with varying doses of corresponding drugs as indicated. A, 96-well plate assay format. B, Final drug serial-concentrations used for each drug.*

Reverse transcription quantitative PCR (RT-qPCR)

RT-qPCR assays were performed as described previously [1-3] with minor modifications. Briefly, mRNA purified directly from cells using the Dynabeads™ mRNA DIRECT™ Purification Kit (61012; ThermoFisher Scientific, UK) were used directly in RT-qPCR reaction containing qPCRBIO SyGrene 1-Step Go Lo-ROX (PB25.31-12; PCRBiosystems, UK) and gene-specific primers for one-step reverse transcription and qPCR to quantify gene expression in the LightCycler 480 qPCR system (Roche, UK) based on our previously published protocols [1, 3, 11] which are MIQE compliant [12]. Briefly, thermocycling begins with 45ºC for 10 mins (for reverse transcription) followed

by 95ºC for 30s prior to 45 cycles of amplification at 95ºC for 1s, 60ºC for 1s, 72ºC for 1s, 78ºC for 1s (data acquisition). A 'touch-down' annealing temperature intervention (66ºC starting temperature with a stepwise reduction of 0.6ºC/cycle; 8 cycles) was introduced prior to the amplification step to maximise primer specificity. Melting analysis (95ºC for 30s, 75ºC for 30s, 75-99ºC at a ramp rate of 0.57ºC/s) was performed at the end of qPCR amplification to validate single product amplification in each well. Relative quantification of mRNA transcripts was calculated based on the second derivative maximum algorithm [13] (Roche). Primer sequences are provided in Additional File 2: Table S2. All target genes were normalised to two stable reference genes validated previously [4] to be amongst the most stable reference genes across a wide variety of primary human epithelial cells, dysplastic and squamous carcinoma cell lines, using the GeNorm algorithm [14]. No template controls (NTCs) were prepared by omitting cells/tissue sample during RNA purification and eluates were used as NTCs for qPCR assays to monitor contamination.

siRNA gene silencing on reversal of chemoresistance

WT and drug-resistant cells were transfected with ON-TARGETplus SMARTpool siRNA (Horizon Dharmacon) for human NEK2 (L-004090-00-0005), DNMT1 (L-004605-00-0005), INHBA (L-011701-00-0005), TOP2A (L-004239-00-0005) or Non-targeting Pool siCTRL (D-001810-10-05) using DharmaFECT 1 transfection reagent (T-2001-02) according to DharmaFECTTM reverse transfection protocol provided by Horizon Dharmacon. Briefly, cells were seeded into 96-well plate (~5,000 cells/well) containing corresponding serial dilution of each drug (at 100x concentration) and siRNA transfection mix (containing siRNA in 0.22 µL/well of DharmaFect 1 transfection reagent) to provide a final siRNA concentration of 55 nM/well. Control cells received the same concentration of transfection reagent but without siRNA added. Details of drug concentrations, 96-well plate assay setup and siRNA reverse transfection protocol are shown in Figure S3 below. Following 3-day incubation, cells were harvested for RT-qPCR to confirm specific siRNA gene knockdown (Additional File 2: Figure S13) and crystal violet cell viability assay to investigate siRNA-induced reversal of chemoresistance (Additional File 2: Figure S14-16).

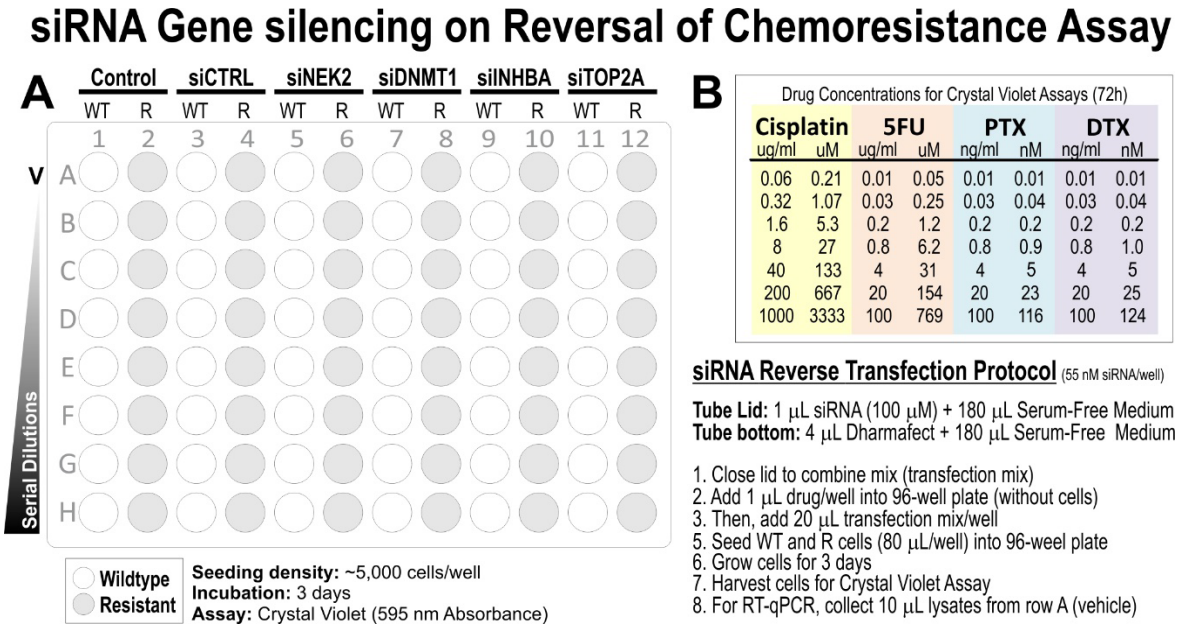


Figure S3. *siRNA gene silencing on reversal of chemoresistance assay. A, 96-well plate format setup with wildtype (WT) and drug-resistant (R) cells were seeded in wells containing corresponding final concentrations of serially diluted drugs (shown in B) and in the presence of corresponding siRNA (55nM final concentration per well) transfection mix.*

Determination of Chemosensitivity (IC_{50})

Chemosensitivity or drug potency on cell viability or gene expression was determined by performing curve-fitting on dose-response data points to calculate the concentration of drug which induced 50% inhibition (IC_{50}) using the four-parameter logistic Hill equation [15]:

$$Y = \frac{A-D}{1+\left(\frac{X}{C}\right)^{B}} + D$$

where Y is the percentage of cell death or gene downregulation, X is the drug concentration, A is the maximal cell density or gene expression, B is the slope factor, C is the IC_{50} and D is the minimal cell density or gene level. Cell viability or gene expression datapoints of dose-response assays were curve-fitted based on the above algorithm using the Quest Graph™ IC_{50} Calculator (AAT Bioquest, Inc, 04 Jul. 2019, https://www.aatbio.com/tools/ic50-calculator).

Pan-cancer and HNSCC transcriptome data mining

Pan-cancer transcriptome datasets were queried in the Oncomine (www.oncomine.org) [16] and Kaplan-Meier Plotter (KM-Plotter.com) [17] databases. The initial differentially expressed gene selection study was performed in Oncomine with the main inclusion criterion that the studies must involve comparison between HNSCC tumour samples with normal tissues. Studies using HNSCC cell lines were excluded. At the time of analysis, there were eight studies eligible for analysis (Additional File 2: Table S1). Differentially expressed genes were ranked according to their median P-values for over-expression and under-expression. Candidate genes were selected based on their top-ranking positions across the eight studies. For differential gene expression of selected candidate genes in HNSCC tumour tissues and matching normal margins, we have used the pan-cancer GEPIA 'Box Plot' tool (GEPIA, http://gepia.cancer-pku.cn/) [18] which is based on transcriptomic data of The Cancer Genome Atlas (TCGA)/The Genotype-Tissue Expression (GTEx) [19]). We have also used GEPIA to survey candidate gene expression profile across 33 human cancer types. For pan-cancer biomarker prognostic survival analysis, hazard ratio (HR) and logrank P values were extracted from each corresponding Kaplan-Meier plots for each cancer type with either single marker or different combinations of 2, 3 or 4 markers (there were a total of 15 unique combinations of 1 to 4 markers studied). The main exclusion criterion was when HR values were associated with logrank P values of <0.05 (Additional File 3).

Drug Library Screen

A total of 537 compounds were obtained from the Drug Synthesis & Chemistry Branch, Developmental Therapeutics Program (DTP) at the National Cancer Institute (National Institute of Health, USA), consisting of 147 approved oncology drugs (AOD IX) and 390 known natural products (Set V) selected from the DTP Open Repository collection of 140,000 compounds. Factors in selection were origin, purity (>90% by ELSD, major peak has correct mass ion), structural diversity and availability of compound. Drug library compounds were each given individual NSC ID number searchable at DTP Chemical database (https://dtp.cancer.gov/dtpstandard/ChemData/index.jsp). Original drug stocks (10 mM in DMSO) were diluted to 0.1 mM (in DMSO) as working stocks arrayed in 384-well reservoir plates for downstream screening using alamarBlue™ Cell Viability Reagent (DAL1025/DAL1100; ThermoFisher Scientific, Paisley, UK) in 384-well format. Cells (4000 cells/well in 384-well plates) were seeded one day before the addition of drugs (final drug concentration at 1 μM) which were incubated for 72h before addition of alamarBlue™ for 24h incubation before measuring fluorescence (excitation 540 nm and emission 590 nm) using a CLARIOstar microplate reader. Candidate drugs were selected based on anti-proliferative effects between wildtype and drug-resistant cells. For drug-gene dose-dependent interaction study, cells (8000 cells/well in 96-well plates) were incubated with candidate drugs (5-fold dilution containing 6 concentrations from 0.32 nM to 1 μM) for 24h prior to harvest for RT-qPCR to investigate their dose-response effects on gene expression. Assay format and detail protocol are shown in Figure S4 below.

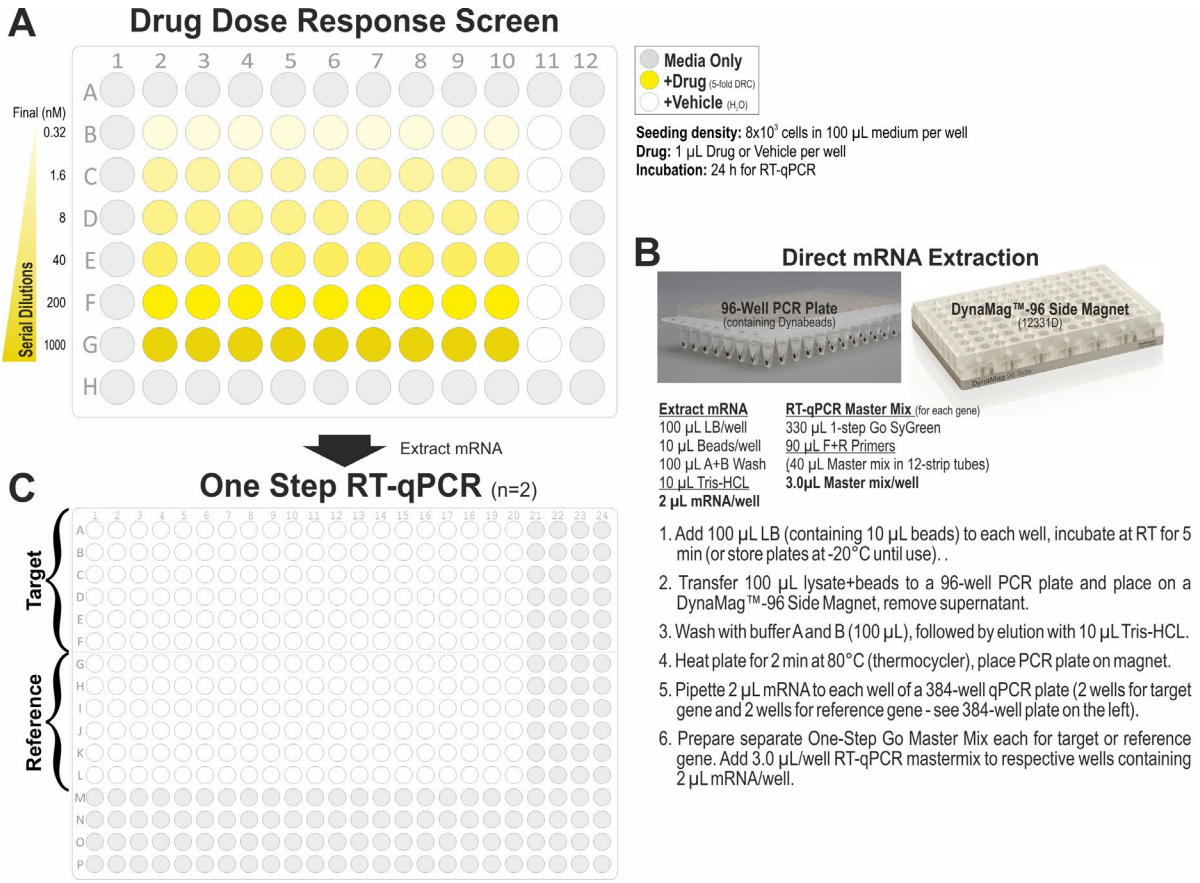


Figure S4. *Drug-gene dose-dependent interaction screening protocol and plate setups. A. Cells (8x10^3 cells/well) were treated for 24h with serial dilution series of nine different drugs as indicated in the diagram. Wells along the edges were not used to eliminate non-specific edge-associated effects. Control cells were treated with equal volume of vehicle (distilled water). B. Direct mRNA extraction method using Dynabeads were performed using a 96-well PCR plate and DynaMag™-96 Side Magnet (12331D) with protocol as illustrated. A total of 60 mRNA samples harvested from panel A (using protocol in panel B) were simultaneously extracted for RT-qPCR in panel C. C. One-step RT-qPCR 384-well plate map for target and reference gene expression quantification in the 60 mRNA samples performed in duplicates.*

Statistical Analysis

Statistical t-tests P values were used for differential analysis between two groups of data. Linear and non-linear regression analyses were used to quantify the relationship between serial drug concentrations and gene expression levels (Additional Data Figure S5-S16). Beeswarm Boxplots were created in R (version 2.13.1; The R Foundation for Statistical Computing) [20].

References

1. Teh MT, Hutchison IL, Costea DE, Neppelberg E, Liavaag PG, Purdie K, Harwood C, Wan H, Odell EW, Hackshaw A, Waseem A: **Exploiting FOXM1-orchestrated molecular network for early squamous cell carcinoma diagnosis and prognosis.** *Int J Cancer* 2013, **132:**2095-2106.
2. Ma H, Dai H, Duan X, Tang Z, Liu R, Sun K, Zhou K, Chen H, Xiang H, Wang J, et al: **Independent evaluation of a FOXM1-based quantitative malignancy diagnostic system (qMIDS) on head and neck squamous cell carcinomas.** *Oncotarget* 2016, **7:**54555-54563.
3. Teh MT, Ma H, Liang YY, Solomon MC, Chaurasia A, Patil R, Tekade SA, Mishra D, Qadir F, Yeung JS, et al: **Molecular Signatures of Tumour and Its Microenvironment for Precise Quantitative Diagnosis of Oral Squamous Cell Carcinoma: An International Multi-Cohort Diagnostic Validation Study.** *Cancers (Basel)* 2022, **14**.
4. Gemenetzidis E, Bose A, Riaz AM, Chaplin T, Young BD, Ali M, Sugden D, Thurlow JK, Cheong SC, Teo SH, et al: **FOXM1 upregulation is an early event in human squamous cell carcinoma and it is enhanced by nicotine during malignant transformation.** *PLoS One* 2009, **4:**e4849.
5. Teh MT, Gemenetzidis E, Chaplin T, Young BD, Philpott MP: **Upregulation of FOXM1 induces genomic instability in human epidermal keratinocytes.** *Mol Cancer* 2010, **9:**45.
6. Kulkarni PS, Sundqvist K, Betsholtz C, Hoglund P, Wiman KG, Zhivotovsky B, Bertolero F, Liu Y, Grafstrom RC: **Characterization of human buccal epithelial cells transfected with the simian virus 40 T-antigen gene.** *Carcinogenesis* 1995, **16:**2515-2521.
7. Rheinwald JG, Beckett MA: **Tumorigenic keratinocyte lines requiring anchorage and fibroblast support cultured from human squamous cell carcinomas.** *Cancer Res* 1981, **41:**1657-1663.
8. Reiss M, Pitman SW, Sartorelli AC: **Modulation of the terminal differentiation of human squamous carcinoma cells in vitro by all-trans-retinoic acid.** *J Natl Cancer Inst* 1985, **74:**1015-1023.
9. Locke M, Heywood M, Fawell S, Mackenzie IC: **Retention of intrinsic stem cell hierarchies in carcinoma-derived cell lines.** *Cancer Res* 2005, **65:**8944-8950.
10. Gemenetzidis E, Elena-Costea D, Parkinson EK, Waseem A, Wan H, Teh MT: **Induction of human epithelial stem/progenitor expansion by FOXM1.** *Cancer Res* 2010, **70:**9515-9526.
11. Qadir F, Aziz MA, Sari CP, Ma H, Dai H, Wang X, Raithatha D, Da Silva LGL, Hussain M, Poorkasreiy SP, et al: **Transcriptome reprogramming by cancer exosomes: identification of novel molecular targets in matrix and immune modulation.** *Mol Cancer* 2018, **17:**97.
12. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, et al: **The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments.** *Clin Chem* 2009, **55:**611-622.
13. Zhao S, Fernald RD: **Comprehensive algorithm for quantitative real-time polymerase chain reaction.** *J Comput Biol* 2005, **12:**1047-1064.
14. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: **Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.** *Genome Biol* 2002, **3:**RESEARCH0034.
15. DeLean A, Munson PJ, Rodbard D: **Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose-response curves.** *Am J Physiol* 1978, **235:**E97-102.
16. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM: **ONCOMINE: a cancer microarray database and integrated data-mining platform.** *Neoplasia* 2004, **6:**1-6.
17. Nagy A, Lanczky A, Menyhart O, Gyorffy B: **Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets.** *Sci Rep* 2018, **8:**9227.
18. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z: **GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses.** *Nucleic Acids Res* 2017, **45:**W98-W102.
19. Consortium GT: **Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans.** *Science* 2015, **348:**648-660.
20. Juul N, Szallasi Z, Eklund AC, Li Q, Burrell RA, Gerlinger M, Valero V, Andreopoulou E, Esteva FJ, Symmans WF, et al: **Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: a retrospective analysis of five clinical trials.** *Lancet Oncol* 2010, **11:**358-365.

Supplementary Information

Table S1 – Summary of HNSCC Microarray Datasets

Anatomical Sites	PMID[a]	GEO[b]	Tumour	Normal	LNM[c]
Nasopharynx	16912175	GSE12452	31	10	
Hypopharynx	14676830	GSE2379	34	4	
Hypopharynx	16205657	GSE1722	6	4	2
Tongue	19138406	GSE13601	37	20	
Tongue	15170515	GSE6631	22	22	
Tongue	18254958	GSE9844	53	22	
Oral Cavity	15381369	GSE3524	16	4	
Oral Cavity & Larynx	14729608	Ginos HN*	41	13	

[a]PubMed ID; [b]Gene Expression Omnibus accession number; [c]Lymph Node Metastasis; *Oncomine Dataset name.

Table S2 - Biomarker Primer Sequences

Gene	Loci	Forward Primer	Reverse Primer	Bp[a]
BIRC5	17q25.3	AGAACTGGCCCTTCTTGGA	ACACTGGGCCAAGTCTGG	104
BUB1B	15q15.1	CAGTCAGACTCTCAGCATCAAGA	CGAGGCAGAAGAACCAGAGA	94
CBX7	22q13.1	CGAGTATCTGGTGAAGTGGAAA	GGGGGTCCAAGATGTGCT	77
CDCA5	11q13.1	AGACATGACTCTCCCTGGAATC	CACTCATCCAGCTCCGTTTT	96
CENPA	2p23.3	CTGCACCCAGTGTTTCTGTC	GAGAGTCCCCGGTATCATCC	63
CLEC3B	3p21.31	AGCAGCATGGAGCTCTGG	CTCCTCAAACATCTTTGTGTTCA	144
CRNN	1q21.3	GAGCAAGAGTTTGCCGATGT	TCCACAGTCCCTGTGTGGT	98
CXCL8	4q13.3	AAGTTTTTGAAGAGGGCTGAGA	TGGCATCTTCACTGATTCTTGGA	74
DNMT1	19p13.2	CGATGTGGCGTCTGTGAG	TGTCCTTGCAGGCTTTACATT	64
DUOX1	15q21.1	GGAGGTTTGGCAAGAAGGT	GCGCTTGAACTGTTGCAC	110
FN1	2q35	AACGTGGGAGAAGCCCTAC	TTGTGTCCTGATCGTTGCAT	113
FOXM1	12p13.33	ACTTTAAGCACATTGCCAAGC	CGTGCAGGGAAAGGTTGT	63
FOXM1B	12p13.33	CCAGGTGTTTAAGCAGCAGA	TCCTCAGCTAGCAGCACCTTG	279
FOXO1	13q14.11	AGGCTGAGGGTTAGTGAGCA	TGAAAGACATCTTTGGACTGCTT	91
FOXO3	6q21	TTCAAGGATAAGGGCGACAG	CGACTATGCAGTGACAGGTTG	77
FOXO4	Xq13.1	ACGAGTGGATGGTCCGTACT	GTGGCGGATCGAGTTCTTC	86
FOXO6	1p34.2	AAGGATAAAGGCGACAGCAA	GTGTGCAGCGACAGGTTG	71
HOXA7	7p15.2	GCCAATTTCCGCATCTACCC	GGTAGCGGTTGAAGTGGAAC	121
INHBA	7p14.1	GCTCAGACAGCTCTTACCACA	AAATTCTCTTTCTGGTCCCCACT	69
IVL	1q21.3	TGCCTGAGCAAGAATGTGAG	TTCCTCATGCTGTTCCCAGT	83
MMP13	11q22.2	TGAGCTGGACTCATTGTCGG	AGGTAGCGCTCTGCAAACTG	94
NEK2	1q32.3	CATTGGCACAGGCTCCTAC	GAGCCATAGTCAAGTTCTTTCCA	90
NR3C1	5q31.3	TCCCTGGTCGAACAGTTTTT	GCTGGATGGAGGAGAGCTTA	77
PLAU	10q22.2	TCACTGGCTTTGGAAAAGAGA	TGGTGACTTCAGAGCCGTAG	126
S100A16	1q21.3	CAAGATCAGCAAGAGCAGCTT	GAGCTTATCCGCAGCCTTC	94
SIRT1	10q21.3	AAATGCTGGCCTAATAGAGTGG	TGGCAAAAACAGATACTGATTACC	75
TOP2A	17q21.2	CAGTGAAGAAGACAGCAGCAAA	AAGCTGGATCCCTTTTAGTTCC	96
VIM	10p13	AGGTGGACCAGCTAACCAAC	TTTCGGCTTCCTCTCTCTGA	123
POLR2A*	17p13.1	TCCGTATTCGCATCATGAAC	TCATCCATCTTGTCCACCAC	73
YAP1*	11q22.1	ACAATGACGACCAATAGCTCAG	CCACTGTCTGTACTCTCATCTCG	77

*Reference genes

Figure S1

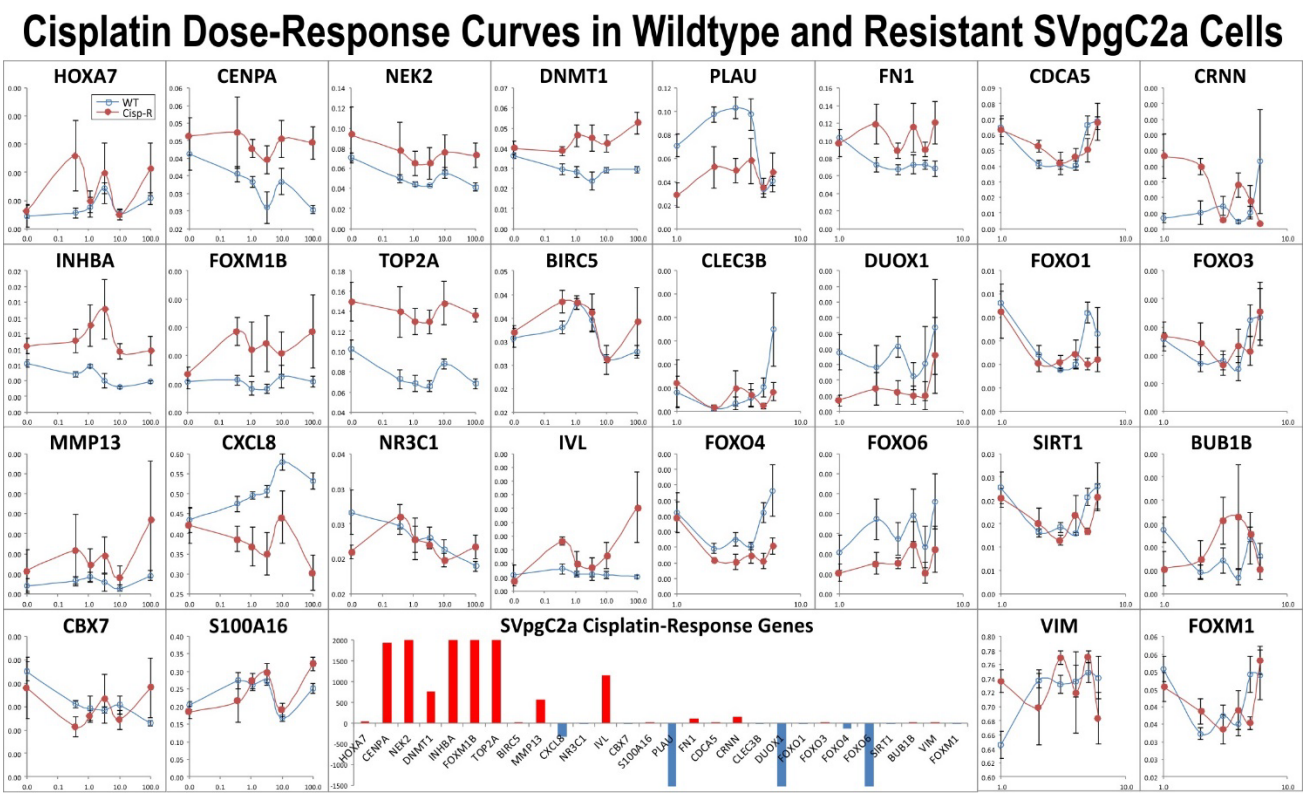


Figure S2

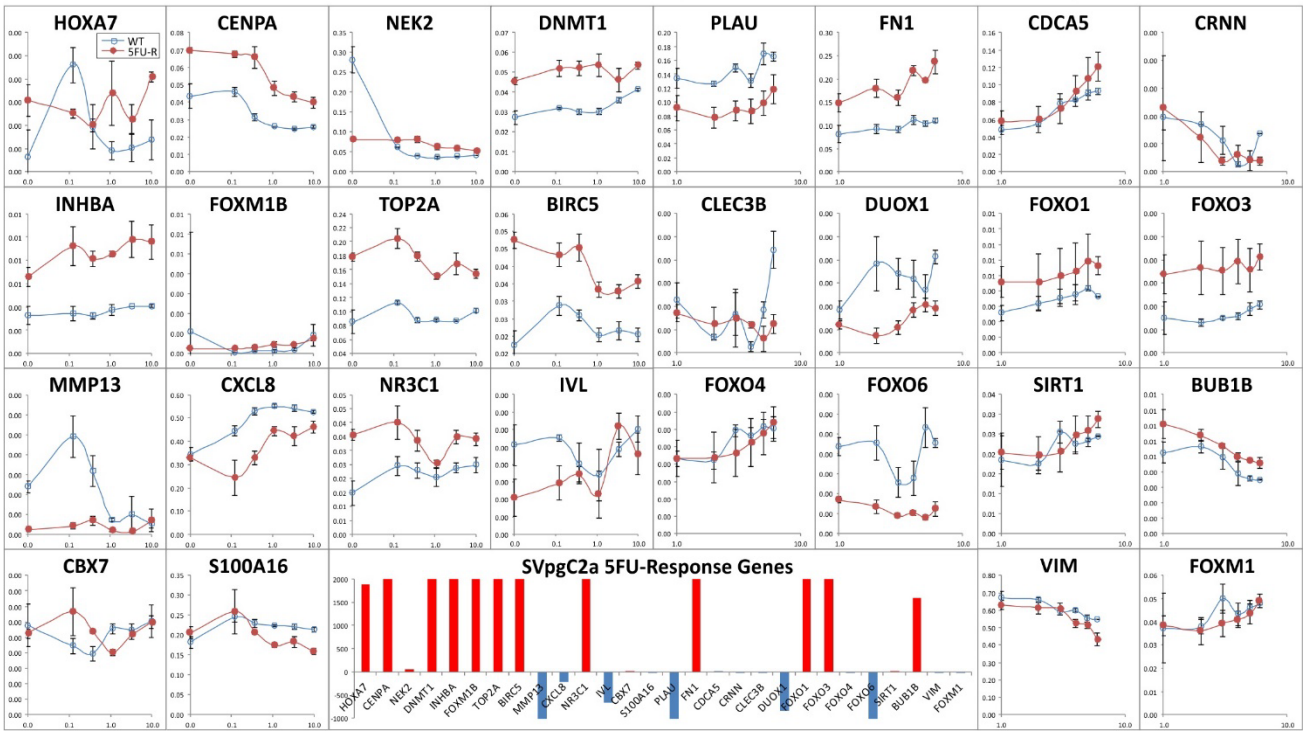


Figure S3

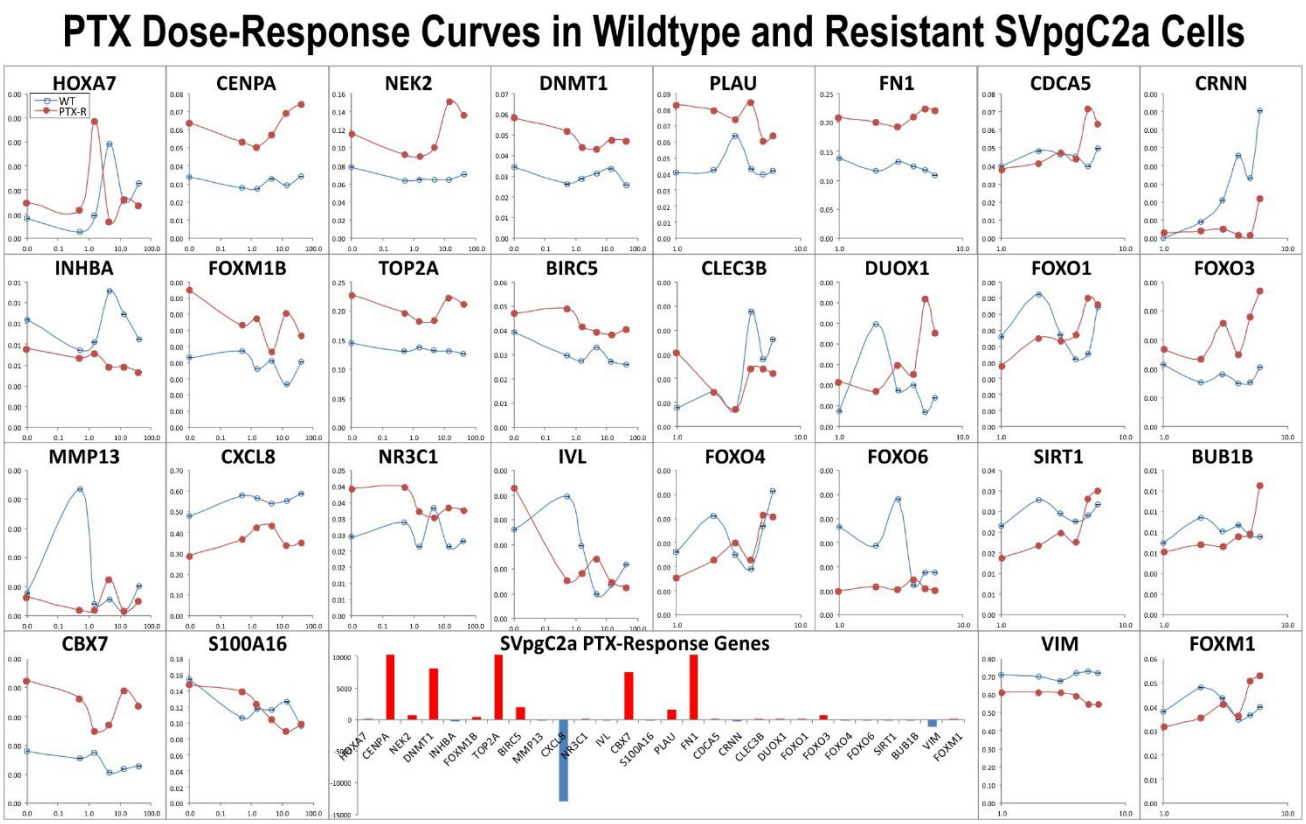


Figure S4

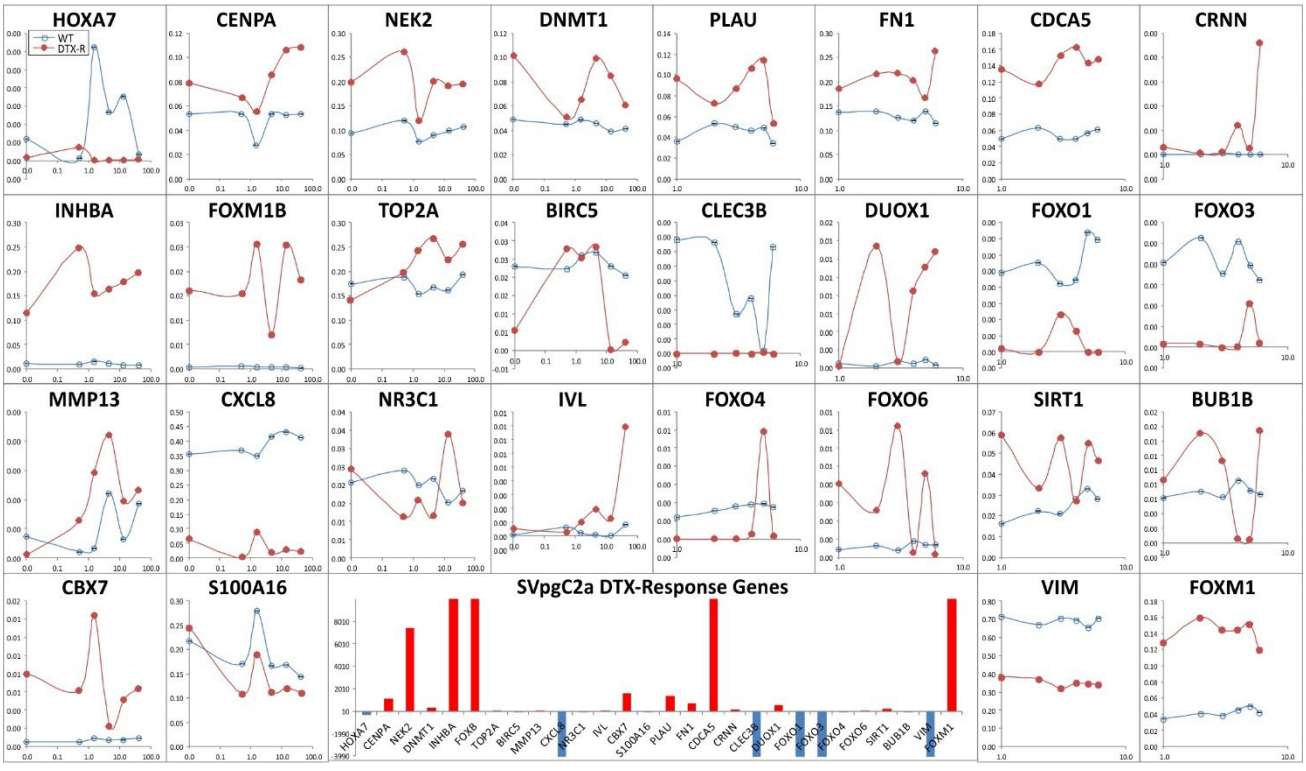


Figure S5

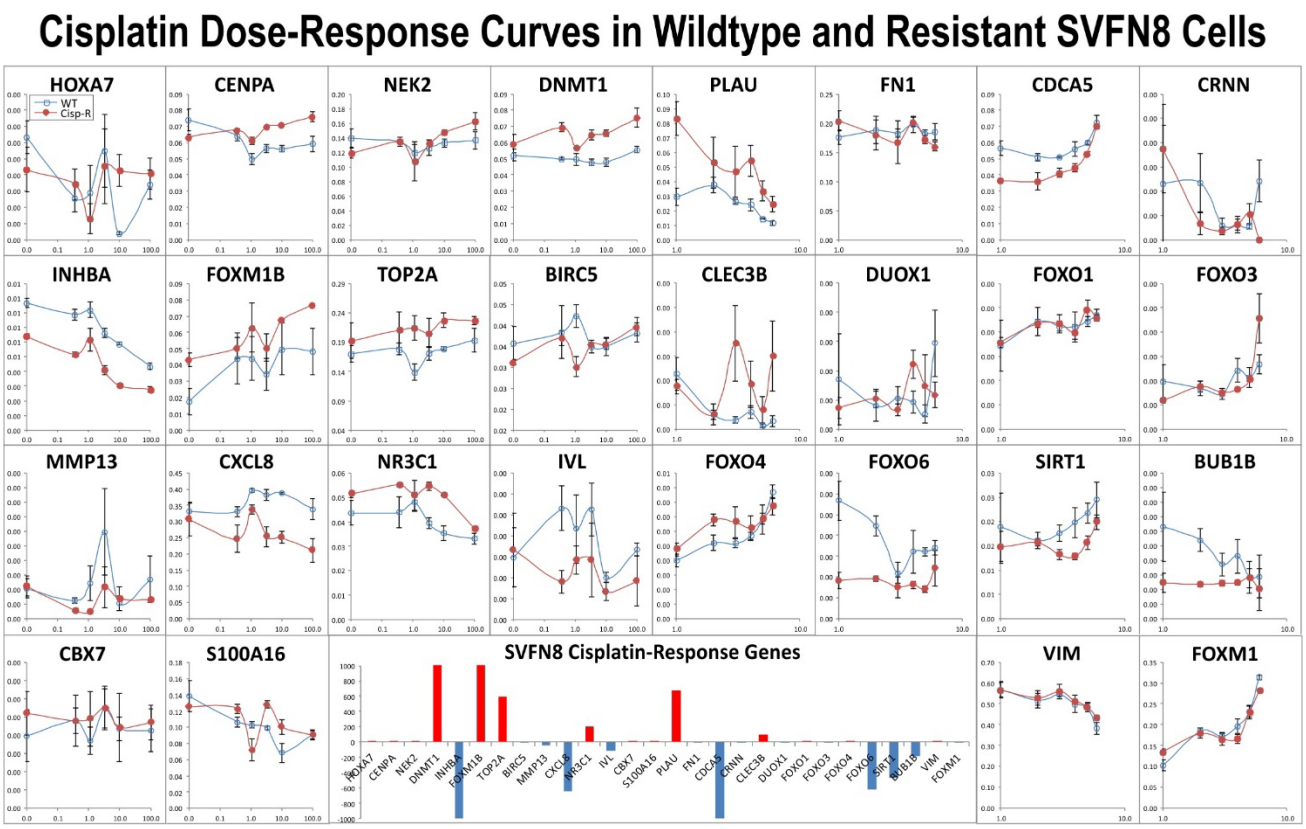


Figure S6

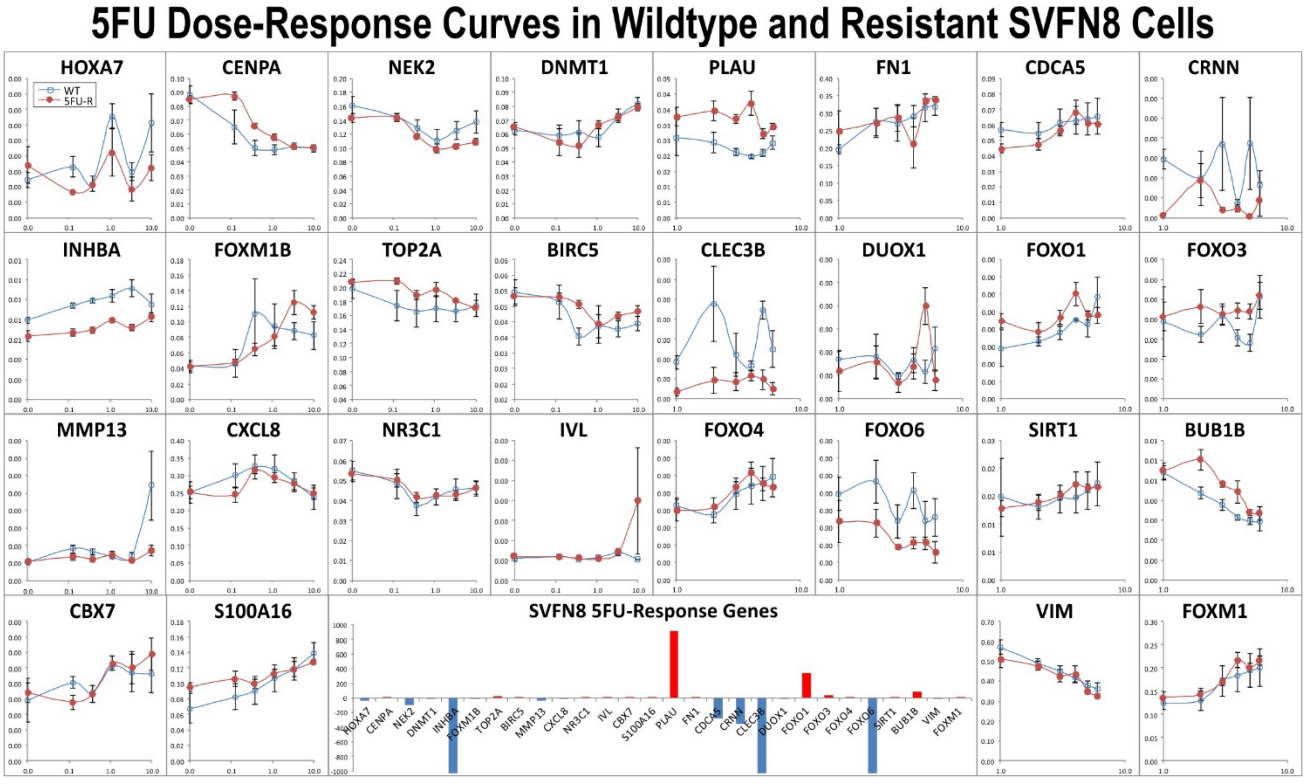


Figure S7

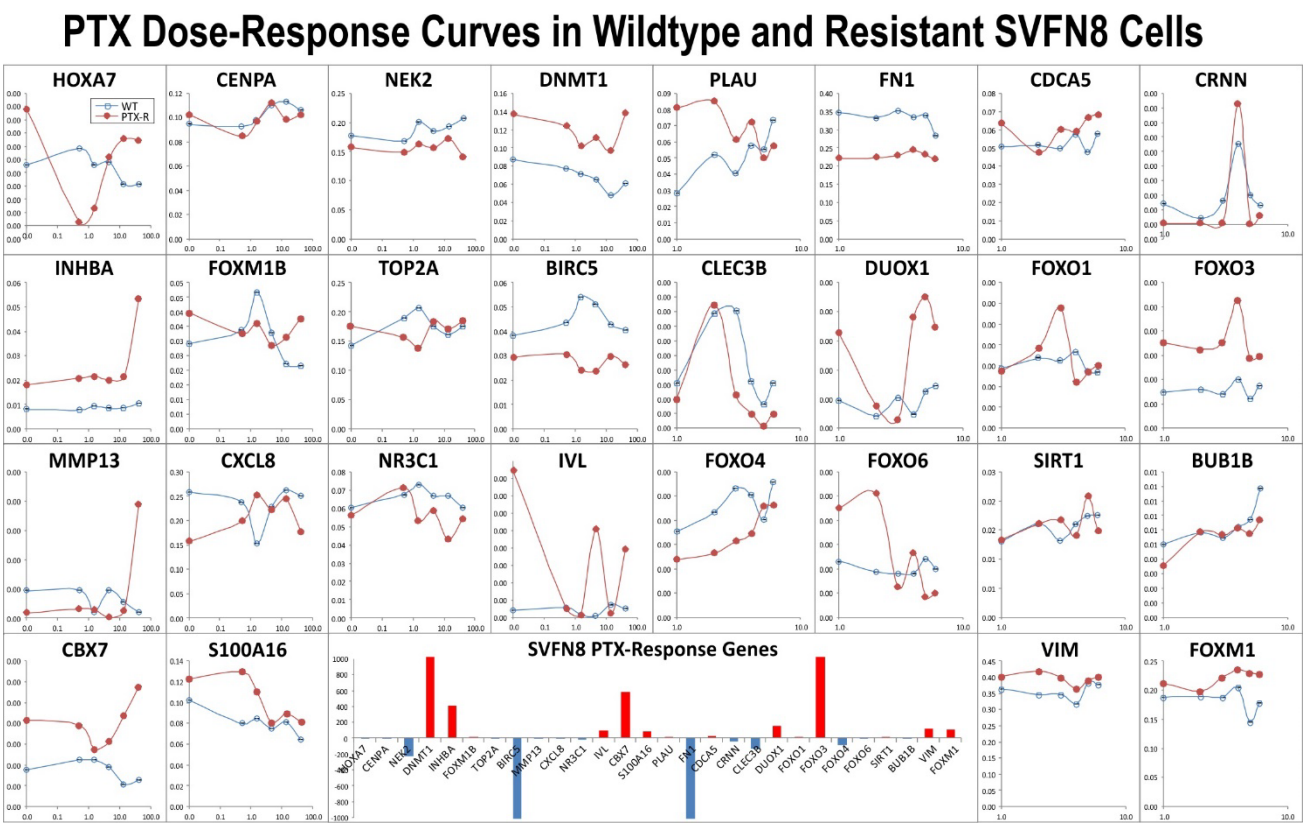


Figure S8

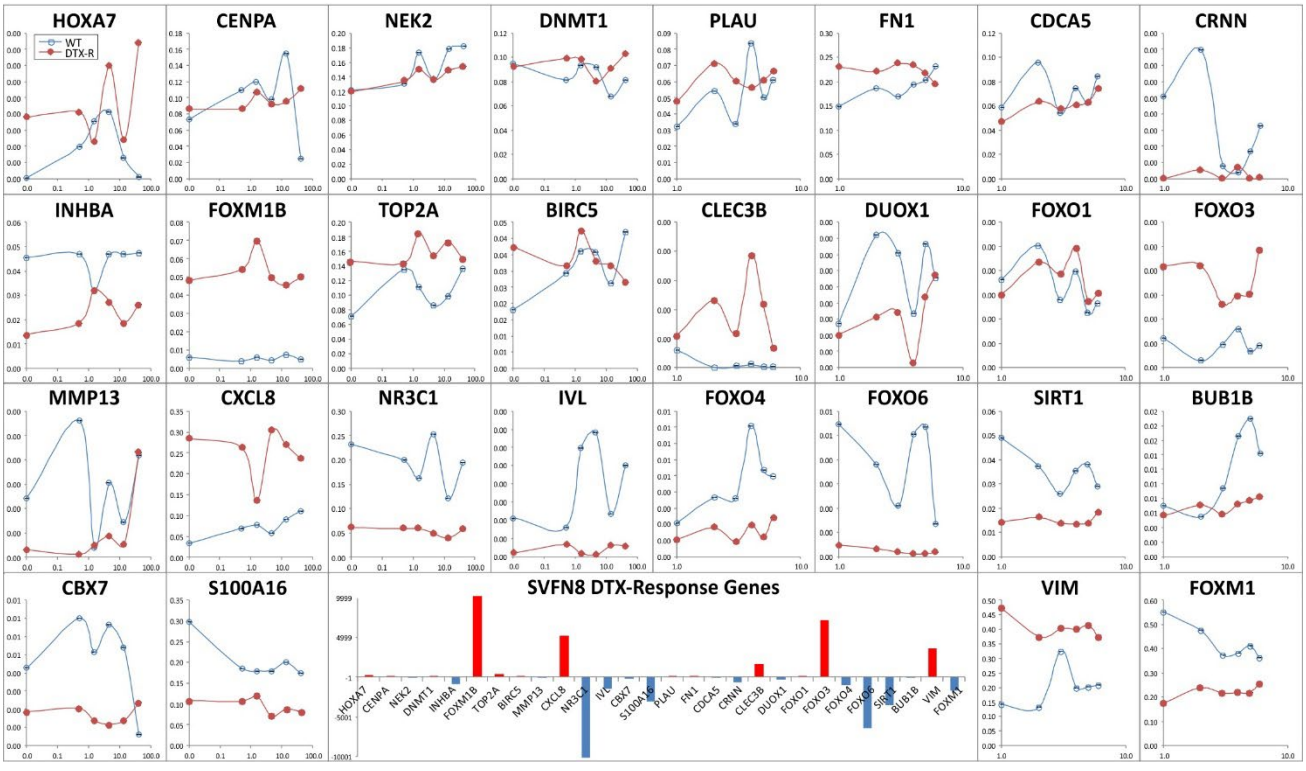


Figure S9

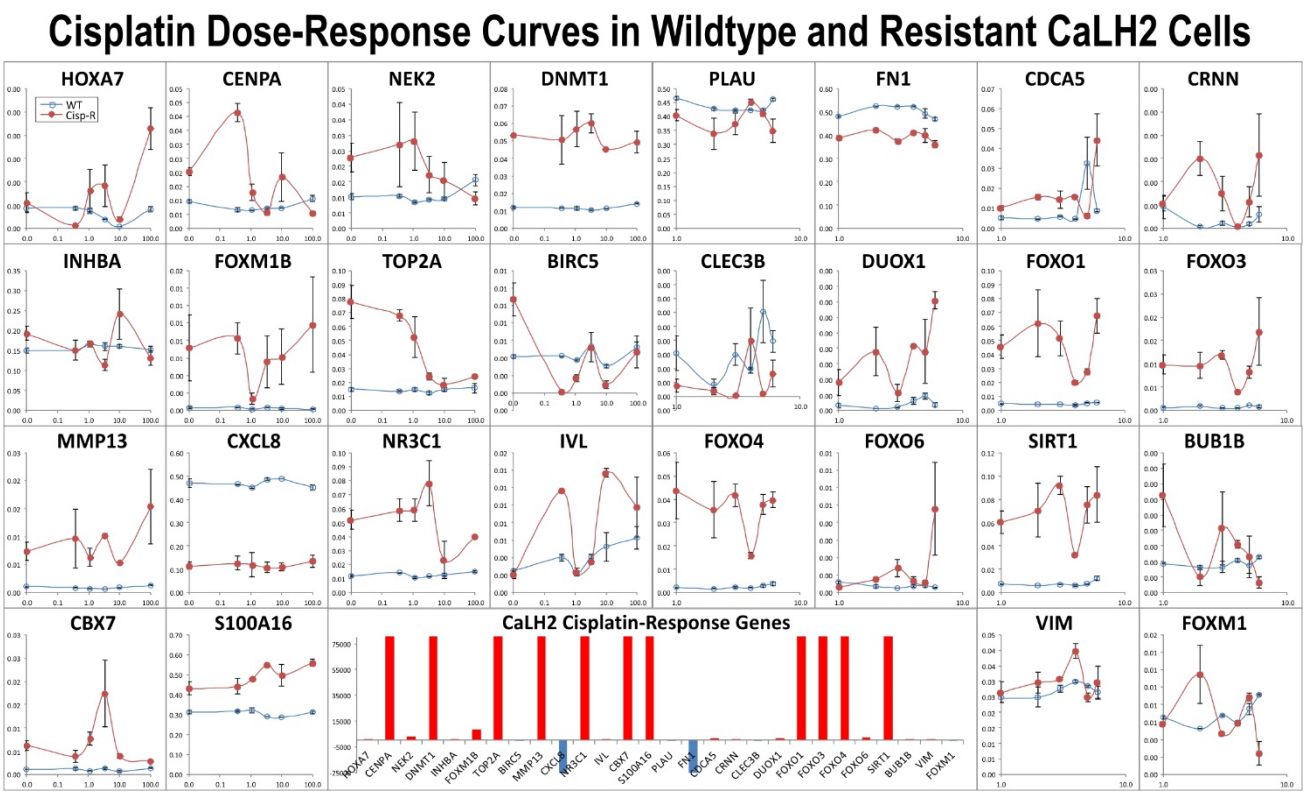


Figure S10

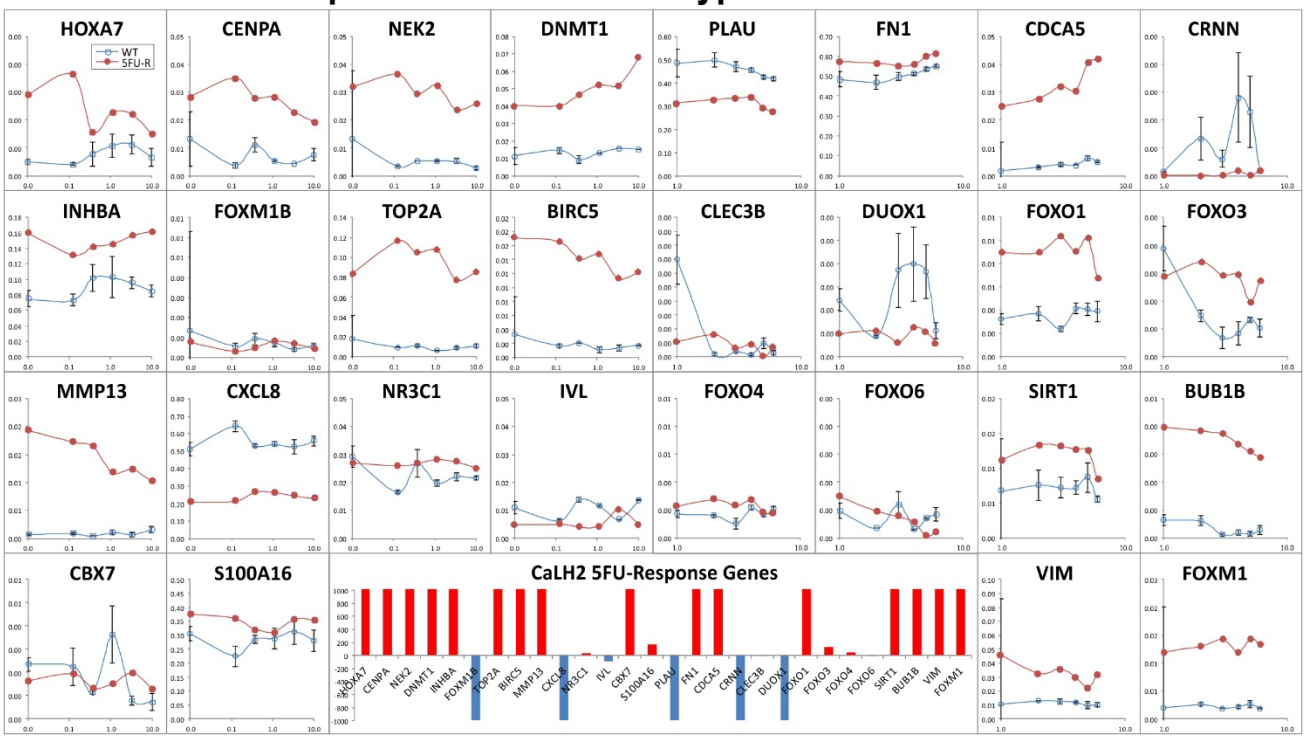


Figure S11

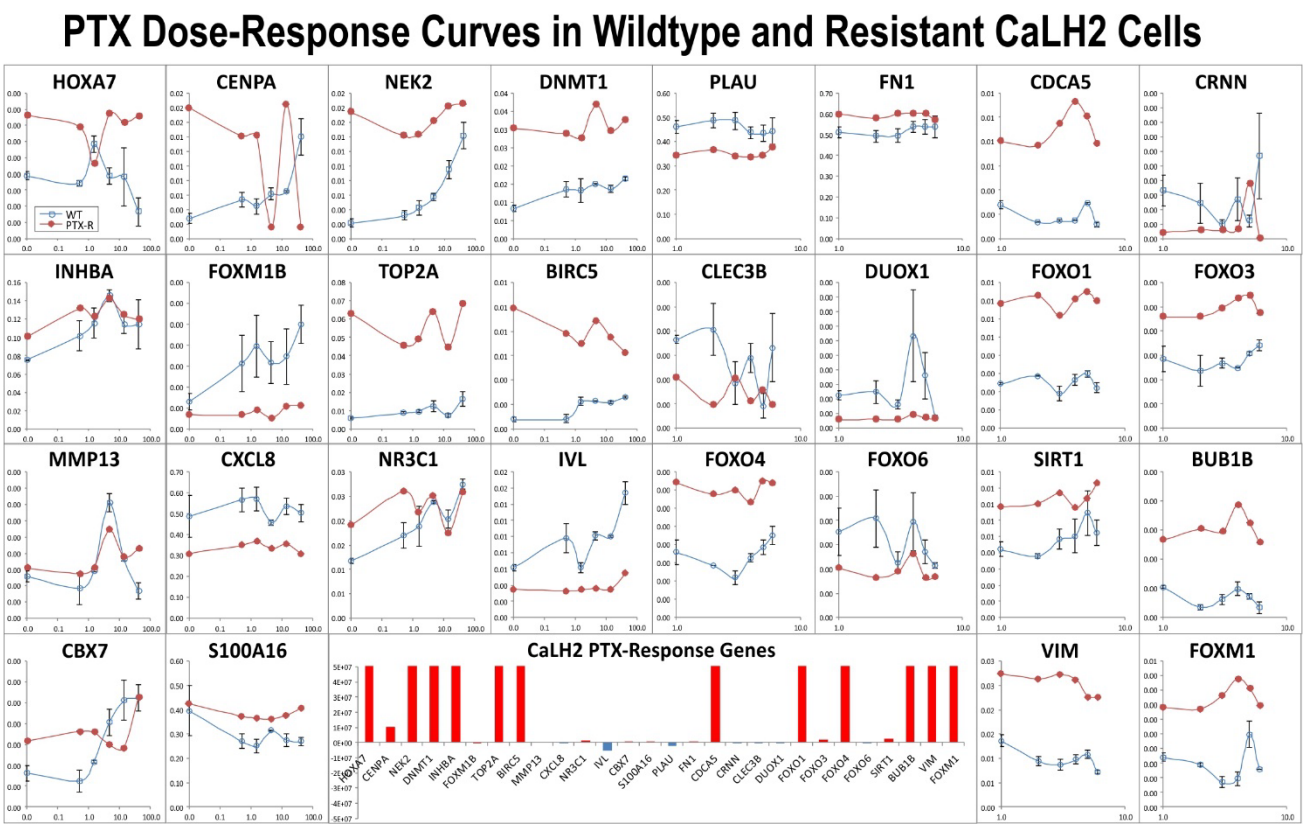


Figure S12

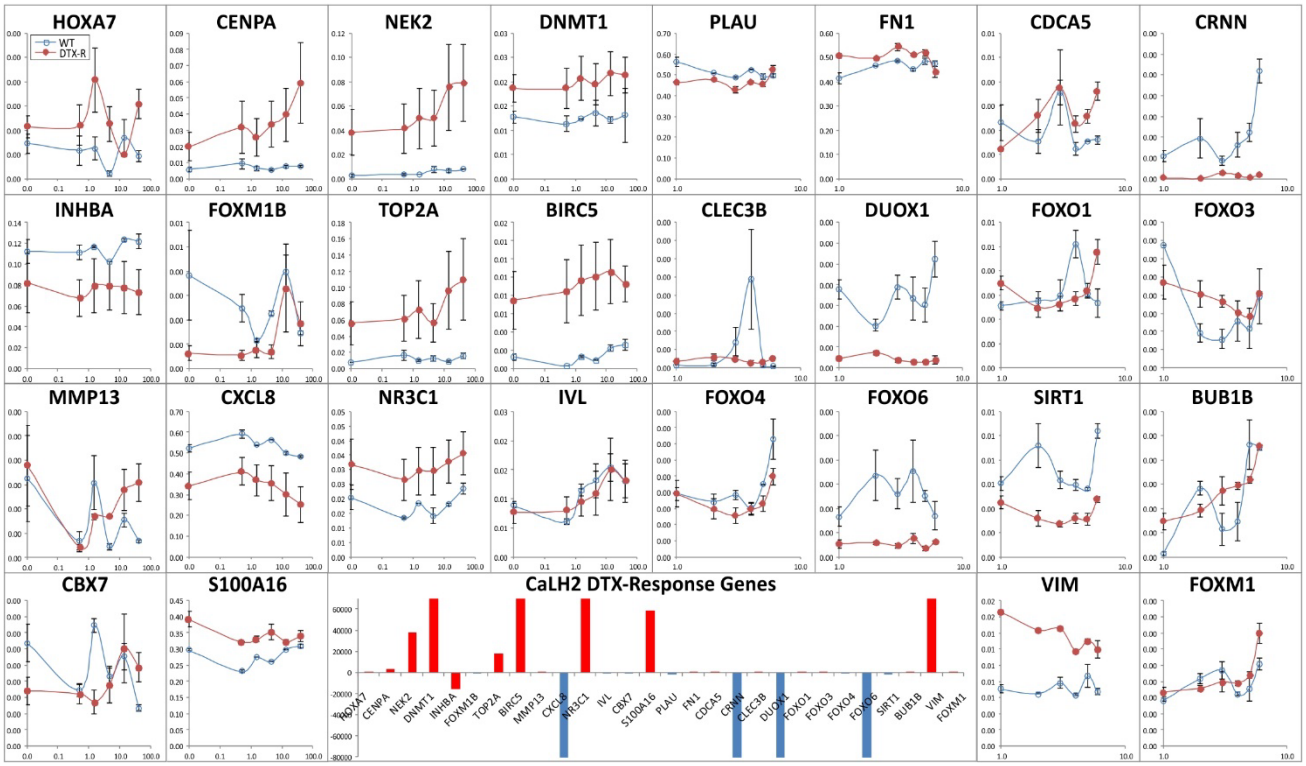


Figure S1-S12. *Differential gene expression analysis on 12 chemoresistant cell strains vs wildtype cells in response to corresponding drugs (cisplatin, 5FU, PTX and DTX). Wildtype (WT; blue line) and drug-resistant (R; red line) cells were treated with serial dilution of corresponding drug for 24h prior to harvest for RT-qPCR to quantify mRNA expression levels of the 28 genes. Assay format and drug concentrations are shown in Additional File 1. Relative gene expression (Y-axis) were plotted against logarithmic drug concentrations (X-axis). Each data point represents a mean of quadruplicates with corresponding SEM error bars. For each gene, the total area between WT and R curves were calculated based on t-test P-values and regression analyses. The most significant up (red bars) or down-regulated (blue bars) genes between WT and R cells are shown as a bar chart in the middle-bottom panel. The significant genes are ranked and tabulated in Figure 1B.*

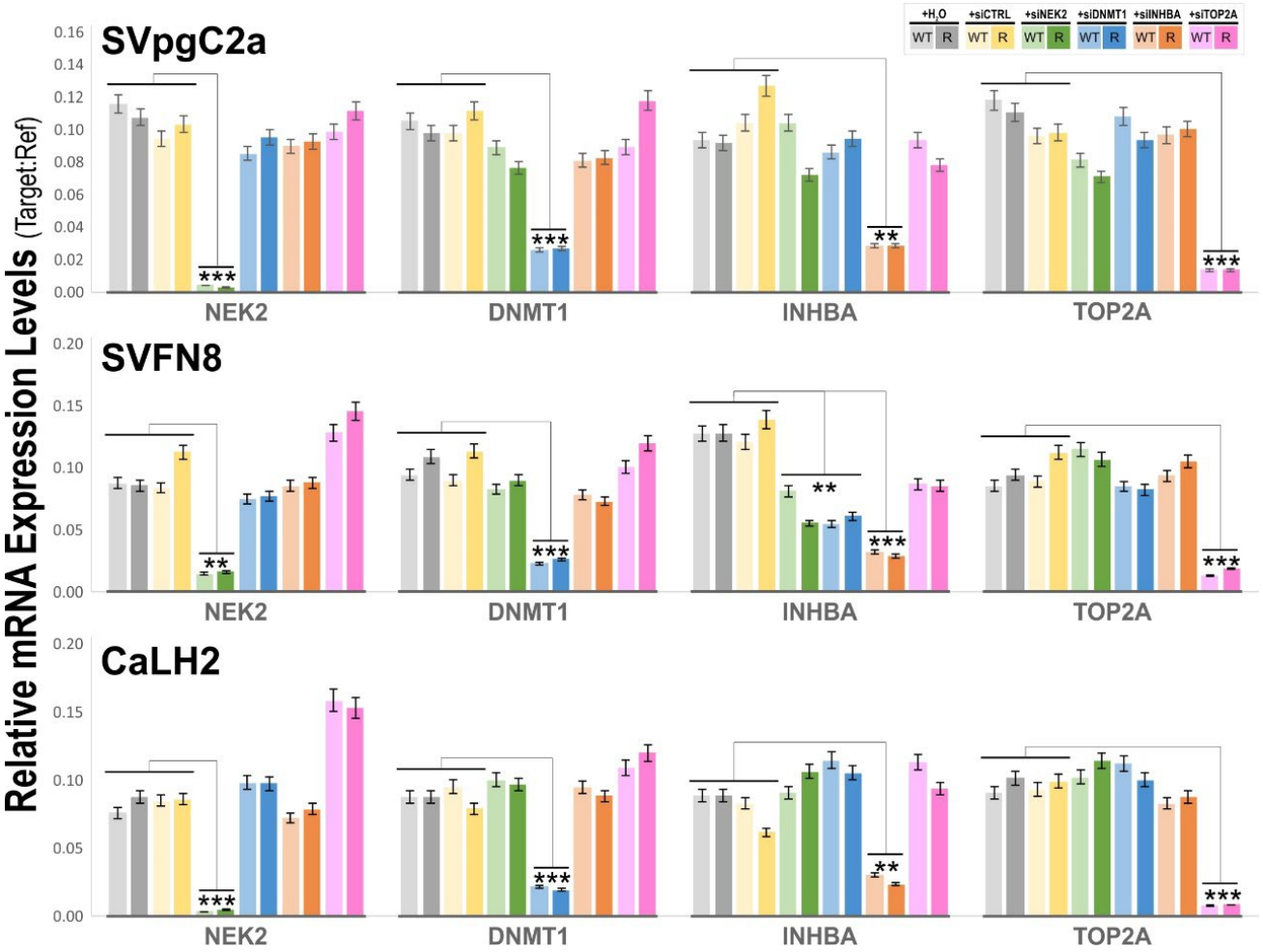


Figure S13. *Validation of gene-specific mRNA knockdown by corresponding siRNA in all 12 cell strains using RT-qPCR. Wildtype (WT) and drug-resistant cells (R) were transfected by either H2O (untransfected control), siCTRL, siNEK2, siDNMT1, siINHBA or siTOP2A as indicated for 3 days followed by RT-qPCR to measure the relative mRNA expression levels of NEK2, DNMT1, INHBA and TOP2A. All 4 siRNAs showed statistically significant gene silencing on corresponding genes (**P<0.01; ***P<0.001) compared to both controls (H2O and siCTRL). None of the siRNAs showed any off-target effects apart from siINHBA on SVFN8 cell lines whereby, in addition to INHBA, NEK2 and DNMT1 mRNA levels were partially but significantly downregulated (**P<0.01). Assay protocol for siRNA transfection and drug concentrations are shown in Additional File 1.*

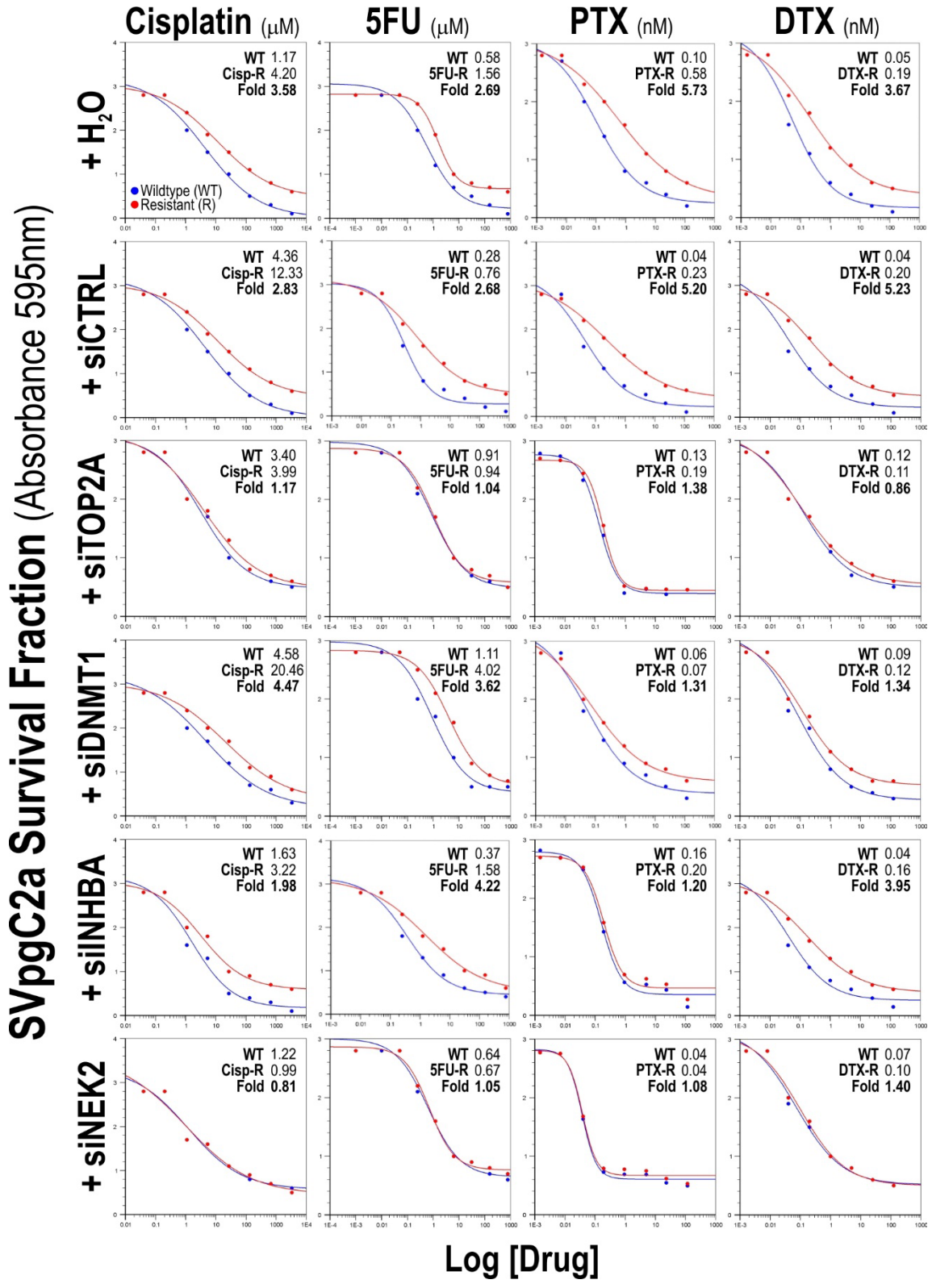


Figure S14. *Effects of siRNA gene silencing of TOP2A, DNMT1, INHBA and NEK2 on chemoresistance in SVpgC2a cells. IC_{50} values of each chemotherapeutic drugs on wildtype (WT) and chemoresistant (R) cells were determined using sigmoid-curve fitting algorithm on data points plotted as logarithmic drug concentrations on the X-axis and survival fraction (Absorbance at 595nm) on the Y-axis. IC_{50} fold differences were calculated between WT and R cells for each treatment as indicated within each graph panel. Assay protocol for siRNA transfection and drug concentrations are shown in Additional File 1.*

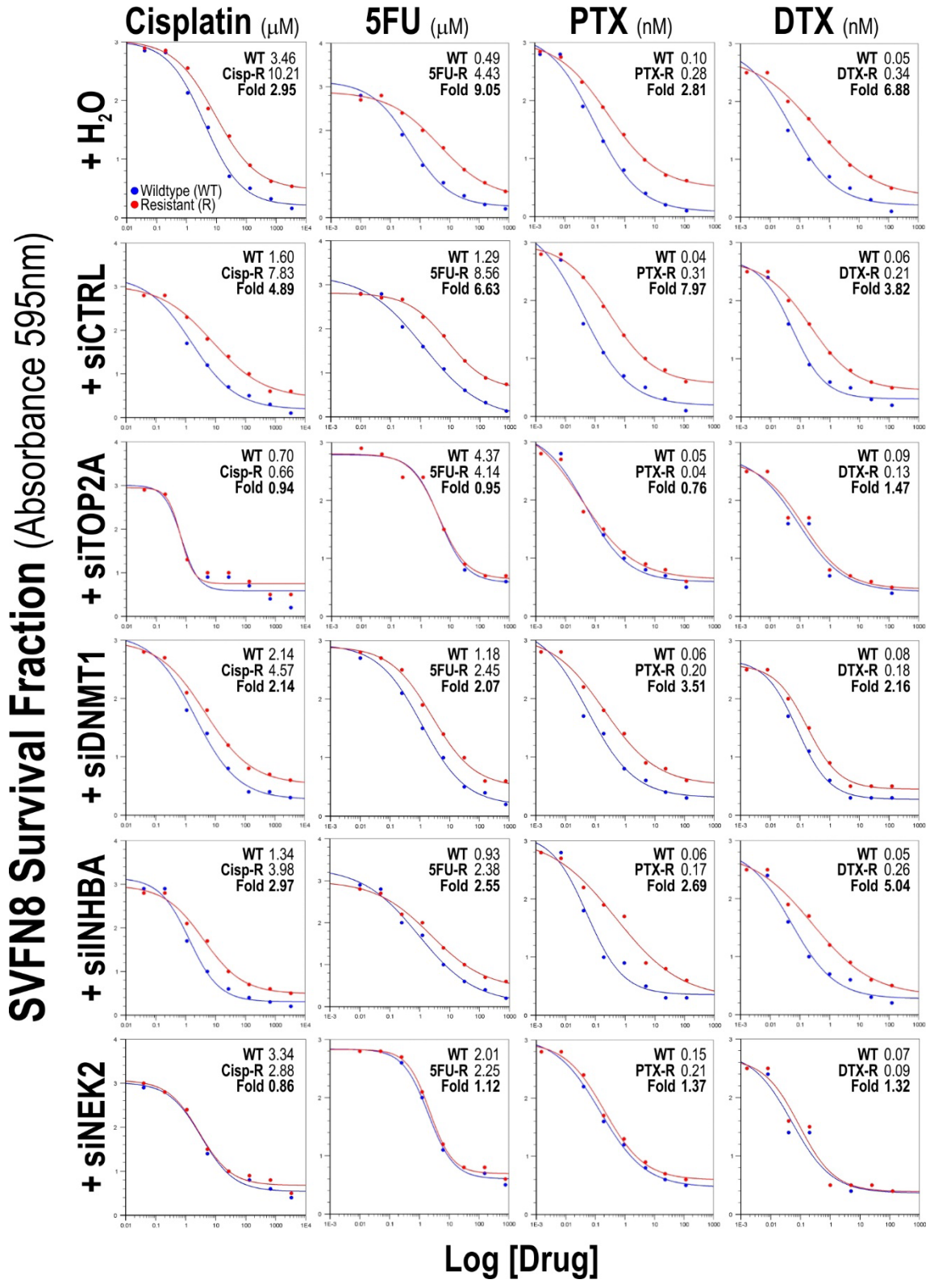


Figure S15. *Effects of siRNA gene silencing of TOP2A, DNMT1, INHBA and NEK2 on chemoresistance in SVFN8 cells. IC_{50} values of each chemotherapeutic drugs on wildtype (WT) and chemoresistant (R) cells were determined using sigmoid-curve fitting algorithm on data points plotted as logarithmic drug concentrations on the X-axis and survival fraction (Absorbance at 595nm) on the Y-axis. IC_{50} fold differences were calculated between WT and R cells for each treatment as indicated within each graph panel. Assay protocol for siRNA transfection and drug concentrations are shown in Additional File 1.*

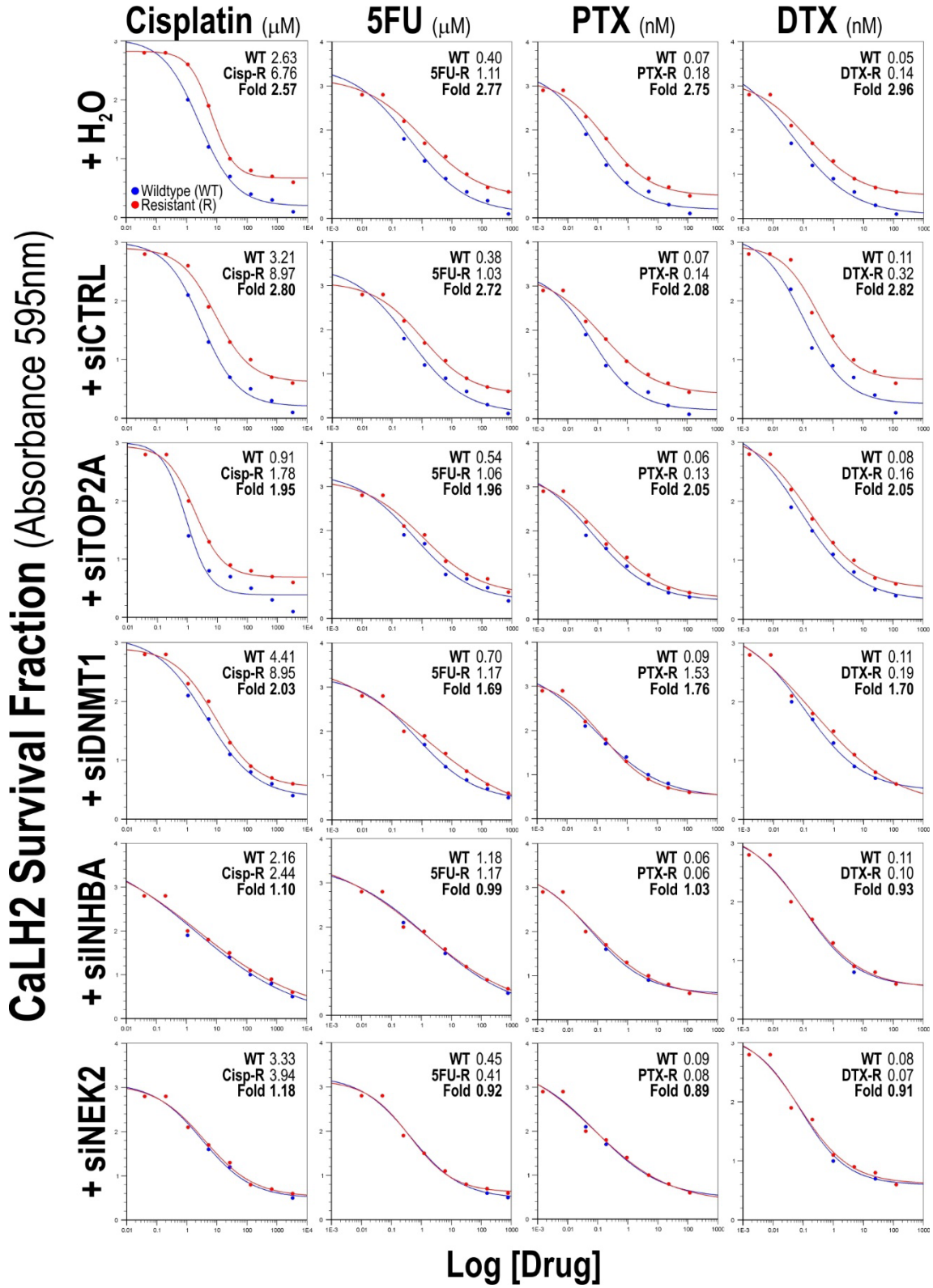
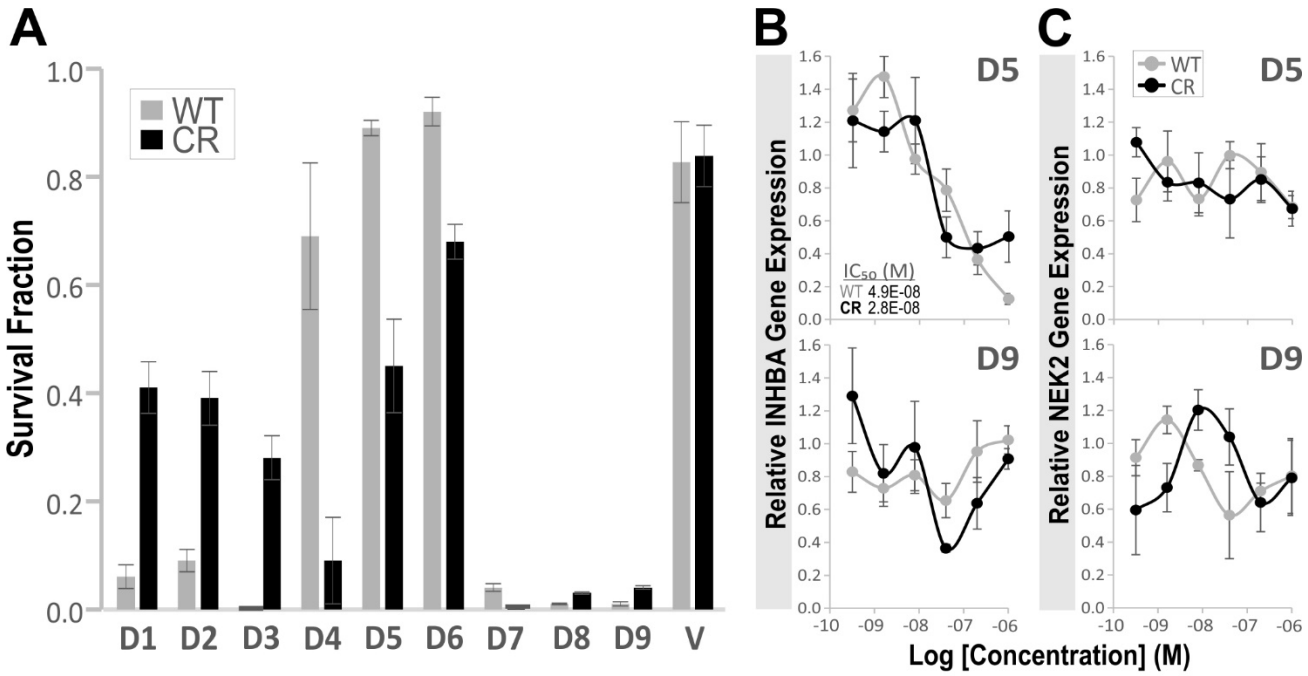


Figure S16. *Effects of siRNA gene silencing of TOP2A, DNMT1, INHBA and NEK2 on chemoresistance in CaLH2 cells. IC_{50} values of each chemotherapeutic drugs on wildtype (WT) and chemoresistant (R) cells were determined using sigmoid-curve fitting algorithm on data points plotted as logarithmic drug concentrations on the X-axis and cell viability (Absorbance at 595nm) on the Y-axis. IC_{50} fold differences were calculated between WT and R cells for each treatment as indicated within each graph panel. Assay protocol for siRNA transfection and drug concentrations are shown in Additional File 1.*



D

	NSC	Chemical name	CAS number	Mol. Formula	Mol. Weight
D1	407806	Digitoxigenin	CAS 143-62-4	C23H34O4	375
D2	761068	Cabozantinib	CAS 849217-68-1	C28H24FN3O5	502
D3	30625	Lasiocarpine	CAS 303-34-4	C21H33NO7	411
D4	294408	Sirodesmin A	CAS 52988-50-8	C20H26N2O8S2	487
D5	44138	1-O-phosphonohexopyranose	CAS 6799-01-5	C6H13O9P	260
D6	121849	CID5351990	N/A	C24H28N2O5	425
D7	758252	Carfilzomib	CAS 868540-17-4	C40H57N5O7	720
D8	269756	Baccharinol	CAS 63783-94-8	C29H38O11	563
D9	526417	Echinomycin A	CAS 512-64-1	C51H64N12O12S2	1101

Figure S17. *Drug library screen to identify drug-gene interactions for counteracting chemoresistance in HNSCC cells.* ***A****, A library of 537 compounds were screened using AlamarBlue cell viability assays on wildtype (WT) and cisplatin-resistant (CR) CaLH2 cell line performed in triplicates. Assay protocol details are shown in Additional File 1. Nine compounds (D1-D9) were selected based on the most significant growth inhibitory effects (t-test, p-value<0.05). Control cells were treated with vehicle (V; 1% DMSO).* ***B-C****, Dose-response assays for D5 and D9 on INHBA (****B****) and NEK2 (****C****) gene expression in WT and CR cells. Each datapoint represents relative gene expression (mean ± SEM) of quadruplicates quantified using RT-qPCR. Drug potencies (IC_{50}) on respective gene inhibition are displayed within each panel. Where dose-response curve fitting could not be performed, no IC50 value is shown.* ***D****, Basic chemical identities of D1-D9 drugs. NSC ID number is searchable at DTP Chemical database (https://dtp.cancer.gov/dtpstandard/ChemData/index.jsp).*

Pan-cancer transcriptome prognostic analysis for 21 cancer types (Data extracted from KMPlot.com)

Hazard Ratios (HR)

	NEK2	TOP2A	INHBA	DNMT1	N+T	N+I	N+D	T+I	T+D	I+D	N+T+I	N+T+D	T+I+D	N+I+D	N+T+I+D
Bladder carcinoma	0.84	1.21	1.58	1.33	0.81	1.62	1.19	1.22	0.83	1.49	1.20	0.82	1.30	1.40	1.32
Breast cancer	1.27	0.80	1.70	0.83	0.83	1.66	0.81	1.37	0.79	1.41	1.38	0.78	1.32	1.58	1.27
Cervical SCC	1.34	0.66	2.94	0.48	0.66	1.66	0.48	1.37	0.57	0.62	1.34	0.56	0.61	0.60	0.59
Esophageal adenocarcinoma	2.76	1.80	2.96	1.62	1.91	1.95	2.40	2.24	1.79	2.49	2.24	1.71	2.02	2.65	1.85
Esophageal SCC	0.40	0.31	0.43	0.19	0.31	0.33	0.17	0.37	0.30	0.30	0.36	0.22	0.27	0.26	0.27
HNSCC	1.44	1.22	1.68	0.69	1.23	1.65	0.74	1.28	0.84	1.17	1.26	0.87	0.85	1.19	0.89
Kidney renal clear cell carcinoma	2.36	2.47	0.82	1.13	2.53	1.31	1.22	1.83	1.79	1.19	1.85	1.84	1.67	1.32	1.77
Kidney renal papillary cell carcinoma	4.81	5.32	3.20	2.63	5.32	5.23	3.42	5.55	5.14	3.32	5.48	5.14	5.98	3.32	6.21
Liver hepatocellular carcinoma	2.14	2.14	1.52	1.92	2.20	1.90	1.92	2.05	2.19	1.90	2.29	2.29	2.09	1.91	2.27
Lung adenocarcinoma	1.85	1.96	1.59	1.28	1.94	1.87	1.51	1.74	1.83	1.40	1.83	1.92	1.64	1.48	1.61
Lung SCC	0.81	0.67	1.47	0.97	0.69	1.53	0.79	1.11	0.72	1.26	1.10	0.72	0.87	1.36	0.88
Ovarian cancer	0.77	1.12	0.80	0.71	0.90	1.16	0.73	1.20	0.82	0.76	1.27	0.86	0.80	0.76	0.79
Pancreatic ductal adenocarcinoma	2.08	2.55	1.65	1.46	2.36	1.75	1.73	2.23	2.61	1.68	2.48	2.36	2.25	1.66	1.91
Pheochromolytoma & Paraganglioma		4.59	0.21	3.00	4.86	1.99	3.04	4.95	8.69	1.69	5.00	8.21	6.60	1.81	6.30
Rectum adenocarcinoma	0.73	0.44	0.56	0.50	0.47	0.60	0.53	0.33	0.26	0.51	0.36	0.34	0.33	0.59	0.33
Sarcoma	1.80	2.20	0.58	1.94	2.07	0.80	1.71	2.03	2.02	1.71	2.01	1.99	2.20	1.89	2.17
Stomach adenocarcinoma	0.74	0.73	1.63	0.66	0.72	1.59	0.69	1.15	0.63	1.35	0.81	0.67	0.72	1.36	0.73
Testicular germ cell tumour	0.14		8.30	3E+08											
Thymoma	0.10	0.12	5E+08	0.14	0.12	0.13	0.14	0.12	0.07	0.07	0.12	0.07	0.07	0.11	0.07
Thyroid carcinoma	3.53	1.90	2.25	0.51	1.97	2.37	0.59	3.56	0.67	2.17	3.74	0.65	2.96	2.18	3.08
Uterine corpus endometrial carcinoma	1.44	1.78	0.72	1.40	1.70	1.38	1.38	1.56	1.58	0.77	1.60	1.50	1.52	0.80	1.45

logrank P values

	NEK2	TOP2A	INHBA	DNMT1	N+T	N+I	N+D	T+I	T+D	I+D	N+T+I	N+T+D	T+I+D	N+I+D	N+T+I+D
Bladder carcinoma	3E-01	3E-01	4E-03	1E-01	2E-01	5E-03	3E-01	2E-01	2E-01	1E-02	3E-01	2E-01	8E-02	4E-02	7E-02
Breast cancer	2E-01	2E-01	2E-03	3E-01	3E-01	3E-03	2E-01	6E-02	2E-01	5E-02	6E-02	1E-01	9E-02	6E-03	2E-01
Cervical SCC	2E-01	9E-02	2E-05	8E-02	1E-01	4E-02	1E-02	2E-01	5E-02	6E-02	3E-01	5E-02	8E-02	3E-02	6E-02
Esophageal adenocarcinoma	3E-02	8E-02	1E-02	1E-01	5E-02	4E-02	9E-03	6E-02	7E-02	1E-02	6E-02	1E-01	3E-01	6E-03	6E-02
Esophageal SCC	3E-02	2E-02	3E-02	3E-03	2E-02	6E-02	2E-03	1E-02	1E-02	3E-03	1E-02	8E-03	1E-03	2E-03	1E-03
HNSCC	1E-02	2E-02	4E-04	8E-03	1E-01	2E-04	3E-02	8E-02	3E-02	3E-01	9E-02	3E-01	3E-01	3E-01	5E-01
Kidney renal clear cell carcinoma	2E-08	3E-09	2E-01	5E-01	7E-10	1E-01	2E-01	9E-05	3E-04	3E-01	7E-05	1E-04	9E-04	7E-02	2E-04
Kidney renal papillary cell carcinoma	2E-08	1E-09	3E-05	1E-03	1E-09	2E-09	2E-05	3E-10	3E-09	3E-05	5E-10	3E-09	6E-11	3E-05	2E-11
Liver hepatocellular carcinoma	9E-06	1E-05	2E-02	4E-04	7E-06	2E-04	2E-04	4E-05	6E-06	2E-04	2E-06	2E-06	2E-05	2E-04	4E-06
Lung adenocarcinoma	5E-05	9E-05	5E-04	1E-01	4E-05	2E-04	7E-03	2E-03	5E-04	4E-02	1E-03	2E-03	1E-03	2E-02	2E-03
Lung SCC	2E-01	4E-03	8E-03	8E-02	8E-03	2E-03	1E-01	5E-01	2E-02	1E-01	5E-01	2E-02	3E-01	5E-02	2E-01
Ovarian cancer	7E-02	4E-01	1E-01	2E-02	5E-01	3E-01	2E-02	2E-01	2E-01	4E-02	1E-01	3E-01	1E-01	4E-02	9E-02
Pancreatic ductal adenocarcinoma	7E-04	4E-05	2E-02	7E-02	8E-05	9E-03	9E-03	3E-03	9E-06	1E-02	1E-03	8E-05	4E-03	2E-02	2E-03
Pheochromolytoma & Paraganglioma	1E-01	6E-02	1E-01	2E-01	5E-02	4E-01	2E-01	5E-02	2E-02	5E-01	5E-02	2E-02	5E-02	5E-01	6E-02
Rectum adenocarcinoma	4E-01	4E-02	2E-01	7E-02	5E-02	2E-01	1E-01	6E-03	9E-03	9E-02	9E-03	3E-02	4E-03	2E-01	4E-03
Sarcoma	5E-03	5E-04	6E-03	3E-03	7E-04	3E-01	9E-02	1E-03	3E-03	2E-03	1E-03	4E-03	2E-04	1E-02	3E-04
Stomach adenocarcinoma	7E-02	9E-02	4E-03	1E-02	9E-02	6E-03	4E-02	4E-01	1E-02	7E-02	3E-01	2E-02	1E-01	6E-02	1E-01
Testicular germ cell tumour	7E-02	1E-01	3E-02	2E-01	9E-02	2E-01	7E-02	1E-01	1E-01	1E-01	8E-02	8E-02	9E-02	1E-01	8E-02
Thymoma	7E-04	2E-04	1E-02	1E-03	2E-03	5E-03	1E-03	2E-03	1E-03	2E-03	2E-03	1E-03	1E-03	1E-03	1E-03
Thyroid carcinoma	2E-02	2E-01	1E-01	2E-01	2E-01	9E-02	3E-01	9E-03	4E-01	1E-01	7E-03	4E-01	8E-02	1E-01	7E-02
Uterine corpus endometrial carcinoma	8E-02	8E-03	1E-01	2E-01	2E-02	1E-01	2E-01	4E-03	4E-02	2E-01	3E-02	7E-02	6E-02	3E-01	9E-03

Select for HR with logrank P<0.05

	NEK2	TOP2A	INHBA	DNMT1	N+T	N+I	N+D	T+I	T+D	I+D	N+T+I	N+T+D	T+I+D	N+I+D	N+T+I+D
Bladder carcinoma			1.58			1.62				1.49				1.40	
Breast cancer			1.70			1.66				1.41				1.58	
Cervical SCC			2.94			1.66	0.48		0.57			0.56		0.60	
Esophageal adenocarcinoma	2.76		2.96			1.95	2.40			2.49				2.65	
Esophageal SCC	0.40	0.31	0.43	0.19	0.31		0.17	0.37	0.30	0.30	0.36	0.22	0.27	0.26	0.27
HNSCC	1.44	1.22	1.68	0.69		1.65	0.74		0.84						
Kidney renal clear cell carcinoma	2.36	2.47			2.53			1.83	1.79		1.85	1.84	1.67		1.77
Kidney renal papillary cell carcinoma	4.81	5.32	3.20	2.63	5.32	5.23	3.42	5.55	5.14	3.32	5.48	5.14	5.98	3.32	6.21
Liver hepatocellular carcinoma	2.14	2.14	1.52	1.92	2.20	1.90	1.92	2.05	2.19	1.90	2.29	2.29	2.09	1.91	2.27
Lung adenocarcinoma	1.85	1.96	1.59		1.94	1.87	1.51	1.74	1.83	1.40	1.83	1.92	1.64	1.48	1.61
Lung SCC		0.67	1.47		0.69	1.53			0.72			0.72			
Ovarian cancer				0.71			0.73			0.76				0.76	
Pancreatic ductal adenocarcinoma	2.08	2.55	1.65		2.36	1.75	1.73	2.23	2.61	1.68	2.48	2.36	2.25	1.66	1.91
Pheochromolytoma & Paraganglioma								4.95	8.69		5.00	8.21	6.60		
Rectum adenocarcinoma		0.44			0.47			0.33	0.26		0.36	0.34	0.33		0.33
Sarcoma	1.80	2.20	0.58	1.94	2.07			2.03	2.02	1.71	2.01	1.99	2.20	1.89	2.17
Stomach adenocarcinoma			1.63	0.66		1.59	0.69		0.63			0.67			
Testicular germ cell tumour			8.30												
Thymoma	0.10	0.12	5E+08	0.14	0.12	0.13	0.14	0.12	0.07	0.07	0.12	0.07	0.07	0.11	0.07
Thyroid carcinoma	3.53							3.56			3.74				
Uterine corpus endometrial carcinoma		1.78			1.70			1.56	1.58		1.60				1.45

Poor vs Good Cancer Prognosis for each marker or markers in combinations

	NEK2	TOP2A	INHBA	DNMT1	N+T	N+I	N+D	T+I	T+D	I+D	N+T+I	N+T+D	T+I+D	N+I+D	N+T+I+D
Poor	9	8	13	3	7	11	5	9	8	8	9	7	7	8	7
Good	2	4	2	5	4	1	6	3	7	3	3	6	3	4	3
Total Cancer	11	12	15	8	11	12	11	12	15	11	12	13	10	12	10
Poor Prognosis	82%	67%	87%	38%	64%	92%	45%	75%	53%	73%	75%	54%	70%	67%	70%
Good Prognosis	18%	33%	13%	63%	36%	8%	55%	25%	47%	27%	25%	46%	30%	33%	30%

Top Prognostic Marker(s) Selection for Each Cancer Type

	NEK2	TOP2A	INHBA	DNMT1	N+T	N+I	N+D	T+I	T+D	I+D	N+T+I	N+T+D	T+I+D	N+I+D	N+T+I+D
Bladder carcinoma	-3	5	427	12	-5	306	4	6	-4	149	5	-5	16	37	19
Breast cancer	8	-4	1000	-3	-3	593	-4	22	-5	30	22	-6	15	277	8
Cervical SCC	6	-7	2E+05	-6	-6	40	-44	6	-12	-11	5	-12	-7	-18	-9
Esophageal adenocarcinoma	95	23	269	12	37	44	282	35	25	208	35	17	7	457	30
Esophageal SCC	-16	-17	-13	-63	-17	-6	-94	-31	-21	-115	-28	-29	-225	-173	-225
HNSCC	111	64	3818	-91	9	6875	-26	16	-30	5	15	-3	-3	4	-2
Kidney renal clear cell carcinoma	2E+08	1E+09	-4	2	3E+09	9	6	2E+04	6885	5	3E+04	2E+04	1815	18	8850
Kidney renal papillary cell carcinoma	2E+08	4E+09	1E+05	2E+03	4E+09	3E+09	2E+05	2E+10	2E+09	1E+05	1E+10	2E+09	9E+10	1E+05	3E+11
Liver hepatocellular carcinoma	2E+05	2E+05	8E+01	5E+03	3E+05	9E+03	1E+04	6E+04	3E+05	9E+03	1E+06	1E+06	1E+05	8E+03	6E+05
Lung adenocarcinoma	4E+04	2E+04	3118	13	5E+04	1E+04	207	870	3978	35	1525	1011	1171	70	700
Lung SCC	-5	-181	188	-12	-90	765	-8	2	-38	13	2	-42	-3	27	-4
Ovarian cancer	-11	3	-6	-47	-2	4	-30	5	-5	-19	10	-3	-7	-17	-9
Pancreatic ductal adenocarcinoma	2811	7E+04	92	21	3E+04	201	204	743	3E+05	129	2557	3E+04	643	111	910
Pheochromolytoma & Paraganglioma	0	75	-2	19	95	5	20	103	457	3	106	342	135	4	113
Rectum adenocarcinoma	-2	-12	-4	-7	-10	-3	-5	-60	-29	-6	-40	-14	-87	-3	-77
Sarcoma	360	4231	-91	647	3136	-3	19	1845	777	713	1546	510	1E+04	172	7483
Stomach adenocarcinoma	-11	-9	398	-51	-8	289	-19	3	-57	19	-3	-28	-7	21	-7
Testicular germ cell tumour	-2	0	307	2E+09	0	0	0	0	0	0	0	0	0	0	0
Thymoma	-139	-632	5E+10	-108	-67	-25	-108	-67	-64	-35	-67	-64	-64	-100	-64
Thyroid carcinoma	177	8	23	-3	10	27	-2	379	-2	16	575	-2	39	16	47
Uterine corpus endometrial carcinoma	17	231	-6	7	113	12	7	422	42	-4	55	23	25	-3	161

Log10 (HR)

	NEK2	TOP2A	INHBA	DNMT1	N+T	N+I	N+D	T+I	T+D	I+D	N+T+I	N+T+D	T+I+D	N+I+D	N+T+I+D
Bladder carcinoma			0.20			0.21				0.17				0.15	
Breast cancer			0.23			0.22				0.15				0.20	
Cervical SCC			0.47			0.22	-0.32		-0.24			-0.25		-0.22	
Esophageal adenocarcinoma	0.44		0.47			0.29	0.38			0.40				0.42	
Esophageal SCC	-0.40	-0.51	-0.37	-0.72	-0.51		-0.77	-0.43	-0.52	-0.52	-0.44	-0.66	-0.57	-0.59	-0.57
HNSCC	0.16	0.09	0.23	-0.16		0.22	-0.13		-0.08						
Kidney renal clear cell carcinoma	0.37	0.39			0.40			0.26	0.25		0.27	0.26	0.22		0.25
Kidney renal papillary cell carcinoma	0.68	0.73	0.51	0.42	0.73	0.72	0.53	0.74	0.71	0.52	0.74	0.71	0.78	0.52	0.79
Liver hepatocellular carcinoma	0.33	0.33	0.18	0.28	0.34	0.28	0.28	0.31	0.34	0.28	0.36	0.36	0.32	0.28	0.36
Lung adenocarcinoma	0.27	0.29	0.20		0.29	0.27	0.18	0.24	0.26	0.15	0.26	0.28	0.21	0.17	0.21
Lung SCC		-0.17	0.17		-0.16	0.18			-0.14			-0.14			
Ovarian cancer				-0.15			-0.14			-0.12				-0.12	
Pancreatic ductal adenocarcinoma	0.32	0.41	0.22		0.37	0.24	0.24	0.35	0.42	0.23	0.39	0.37	0.35	0.22	0.28
Pheochromolytoma & Paraganglioma								0.69	0.94		0.70	0.91	0.82		
Rectum adenocarcinoma		-0.36			-0.33			-0.48	-0.59		-0.44	-0.47	-0.48		-0.48
Sarcoma	0.26	0.34	-0.24	0.29	0.32			0.31	0.31	0.23	0.30	0.30	0.34	0.28	0.34
Stomach adenocarcinoma			0.21	-0.18		0.20	-0.16		-0.20			-0.17			
Testicular germ cell tumour			0.92												
Thymoma	-1.00	-0.92	8.66	-0.85	-0.92	-0.89	-0.85	-0.92	-1.15	-1.15	-0.92	-1.15	-1.15	-0.96	-1.15
Thyroid carcinoma	0.55							0.55			0.57				
Uterine corpus endometrial carcinoma		0.25			0.23			0.19	0.20		0.20				0.16

Best Prognostic Marker(s) for Each Cancery Type

		HR*	logrank P
Bladder carcinoma	INHBA	1.58	4E-03
Breast cancer	INHBA	1.70	2E-03
Cervical SCC	INHBA	2.94	2E-05
Esophageal adenocarcinoma	N+I+D	2.65	6E-03
Esophageal SCC	T+I+D	0.27	1E-03
HNSCC	N+I	1.65	2E-04
Kidney renal clear cell carcinoma	N+T	2.53	7E-10
Kidney renal papillary cell carcinoma	N+T+I+D	6.21	2E-11
Liver hepatocellular carcinoma	N+T+I	2.29	2E-06
Lung adenocarcinoma	N+T	1.94	4E-05
Lung SCC	N+I	1.53	2E-03
Ovarian cancer	DNMT1	0.71	2E-02
Pancreatic ductal adenocarcinoma	T+D	2.61	9E-06
Pheochromolytoma & Paraganglioma	T+D	8.69	2E-02
Rectum adenocarcinoma	T+I+D	0.33	4E-03
Sarcoma	T+I+D	2.20	2E-04
Stomach adenocarcinoma	INHBA	1.63	4E-03
Testicular germ cell tumour	INHBA	8.30	3E-02
Thymoma	INHBA	5E+08	2E-04
Thyroid carcinoma	N+T+I	3.74	7E-03
Uterine corpus endometrial carcinoma	T+I+D	1.56	4E-03

*Red: poor prognosis; Green: good prognosis

Select for HR with logrank P<0.05

	NEK2	TOP2A	INHBA	DNMT1
Bladder carcinoma			1.58	
Breast cancer			1.70	
Cervical SCC			2.94	
Esophageal adenocarcinoma	2.76		2.96	
Esophageal SCC	0.40	0.31	0.43	0.19
HNSCC	1.44	1.22	1.68	0.69
Kidney renal clear cell carcinoma	2.36	2.47		
Kidney renal papillary cell carcinoma	4.81	5.32	3.20	2.63
Liver hepatocellular carcinoma	2.14	2.14	1.52	1.92
Lung adenocarcinoma	1.85	1.96	1.59	
Lung SCC		0.67	1.47	
Ovarian cancer				0.71
Pancreatic ductal adenocarcinoma	2.08	2.55	1.65	
Rectum adenocarcinoma		0.44		
Sarcoma	1.80	2.20	0.58	1.94
Stomach adenocarcinoma			1.63	0.66
Testicular germ cell tumour			8.30	
Thymoma	0.10	0.12	######	0.14
Thyroid carcinoma	3.53			
Uterine corpus endometrial carcinoma		1.78		

Log10

	NEK2	TOP2A	INHBA	DNMT1
Bladder carcinoma	0.44	-0.51	0.20	-0.72
Breast cancer	-0.40	0.09	0.23	-0.16
Cervical SCC	0.16	0.39	0.47	0.42
Esophageal adenocarcinoma	0.37	0.73	0.47	0.28
Esophageal SCC	0.68	0.33	-0.37	-0.15
HNSCC	0.33	0.29	0.23	0.29
Kidney renal clear cell carcinoma	0.27	-0.17	0.51	-0.18
Kidney renal papillary cell carcinoma	0.32	0.41	0.18	-0.85
Liver hepatocellular carcinoma	0.26	-0.36	0.20	
Lung adenocarcinoma	-1.00	0.34	0.17	
Lung SCC	0.55	-0.92	0.22	
Ovarian cancer		0.25	-0.24	
Pancreatic ductal adenocarcinoma			0.21	
Rectum adenocarcinoma			0.92	
Sarcoma			8.66	

Log HR Rank sorted for each gene

		NEK2
Kidney renal papillary cell carcinoma	RPCC	0.68
Thyroid carcinoma	TC	0.55
Esophageal adenocarcinoma	EA	0.44
Kidney renal clear cell carcinoma	RCCC	0.37
Liver hepatocellular carcinoma	HC	0.33
Pancreatic ductal adenocarcinoma	PDA	0.32
Lung adenocarcinoma	LA	0.27
Sarcoma	S	0.26
HNSCC	HNSCC	0.16
Esophageal SCC	ESCC	-0.40
Thymoma	T	-1.00

		TOP2A
Kidney renal papillary cell carcinoma	RPCC	0.73
Pancreatic ductal adenocarcinoma	PDA	0.41
Kidney renal clear cell carcinoma	RCCC	0.39
Sarcoma	S	0.34
Liver hepatocellular carcinoma	HC	0.33
Lung adenocarcinoma	LA	0.29
Uterine corpus endometrial carcinoma	UCEC	0.25
HNSCC	HNSCC	0.09
Lung SCC	LSCC	-0.17
Rectum adenocarcinoma	RA	-0.36
Esophageal SCC	ESCC	-0.51
Thymoma	T	-0.92

		INHBA
Thymoma	T	8.66
Testicular germ cell tumour	TGCT	0.92
Kidney renal papillary cell carcinoma	RPCC	0.51
Esophageal adenocarcinoma	EA	0.47
Cervical SCC	CSCC	0.47
Breast cancer	BrC	0.23
HNSCC	HNSCC	0.23
Pancreatic ductal adenocarcinoma	PDA	0.22
Stomach adenocarcinoma	SA	0.21
Lung adenocarcinoma	LA	0.20
Bladder carcinoma	BldC	0.20
Liver hepatocellular carcinoma	LHC	0.18
Lung SCC	LSCC	0.17
Sarcoma	SA	-0.24
Esophageal SCC	ESCC	-0.37

		DNMT1
Kidney renal papillary cell carcinoma	RPCC	0.42
Sarcoma	S	0.29
Liver hepatocellular carcinoma	HC	0.28
Ovarian cancer	OC	-0.15
HNSCC	HNSCC	-0.16
Stomach adenocarcinoma	SA	-0.18
Esophageal SCC	ESCC	-0.72
Thymoma	T	-0.85

	NEK2	%
Poor Prognosis	9	81.8%
Good Prognosis	2	18.2%
Total Cancer	11	100%

	TOP2A	%
Poor Prognosis	8	66.7%
Good Prognosis	4	33.3%
Total Cancer	12	100%

	INHBA	%
Poor Prognosis	13	86.7%
Good Prognosis	2	13.3%
Total Cancer	15	100%

	DNMT1	%
Poor Prognosis	3	37.5%
Good Prognosis	5	62.5%
Total Cancer	8	100%

BldC	Bladder carcinoma
BrC	Breast cancer
CSCC	Cervical SCC
EA	Esophageal adenocarcinoma
ESCC	Esophageal SCC
HNSCC	HNSCC
RCCC	Kidney renal clear cell carcinoma
RPCC	Kidney renal papillary cell carcinoma
HC	Liver hepatocellular carcinoma
LA	Lung adenocarcinoma
LSCC	Lung SCC
OC	Ovarian cancer
PDA	Pancreatic ductal adenocarcinoma
RA	Rectum adenocarcinoma
S	Sarcoma
SA	Stomach adenocarcinoma
TGCT	Testicular germ cell tumour
T	Thymoma
TC	Thyroid carcinoma
UCEC	Uterine corpus endometrial carcinoma