Elucidating Compound Mechanism of Action and Polypharmacology with a Large-scale Perturbational Profile Compendium.

Lucas ZhongMing Hu[1], Eugene Douglass[1,3], Mikko Turunen[1], Sergey Pampou[1,3,4], Adina Grunn[1], Ron Realubit[1,,3,4], Albert A. Antolin[5], Alexander L.E. Wang[1], Hai Li[1,3,4], Prem Subramaniam[1], Prabhjot S Mundi[1], Charles Karan[1,3,4], Mariano Alvarez[6], Andrea Califano[1,3,4,8,9,9]

[1] Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, USA 10032.

[2] Pharmaceutical and Biomedical Sciences, University of Georgia, 250 W. Green Street, Athens, GA 30602

[3] Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, 1130 Saint Nicholas Ave, New York, NY USA 10032.

[4] JP Sulzberger Columbia Genome Center, Columbia University Medical Center, New York, NY 10032, USA

[5] Centre for Cancer Discovery, The Division of Cancer therapeutics, The Institute of Cancer Research, London, UK

[6] DarwinHealth, Inc. New York, NY 10018

[7] Department of Medicine, Columbia University Irving Medical Center, 630 W 168th Street, New York, NY USA 10032.

[8] Department of Biochemistry & Molecular Biophysics, Columbia University Irving Medical Center, 701 W 168th Street, New York, NY USA 10032.

[9] Department of Biomedical Informatics, Columbia University Irving Medical Center, 622 W 168th Street, New York, NY USA 10032.

Corresponding Author: Andrea Califano (ac2248@cumc.columbia.edu)

Authors Disclosures: A.C. is founder, equity holder, consultant, and director of DarwinHealth Inc., a company that has licensed some of the algorithms used in this manuscript from Columbia University. M.J.A. is Chief Scientific Officer and equity holder at DarwinHealth, Inc., a company that has licensed some of the algorithms used in this manuscript from Columbia University. Charles Karan is a DarwinHealth consultant and shareholder.

Abstract (189 words):

The Mechanism of Action (MoA) of a drug is generally represented as a small, non-tissue-specific repertoire of high-affinity binding targets. Yet, drug activity and polypharmacology are increasingly associated with a broad range of off-target and tissue-specific effector proteins. To address this challenge, we have leveraged a microfluidics-based Plate-Seq technology to survey drug perturbational profiles representing >700 FDA-approved and experimental oncology drugs, in cell lines selected as high-fidelity models of 23 aggressive tumor subtypes. Built on this dataset, we implemented an efficient computational framework to define a tissue-specific protein activity landscape of these drugs and reported almost 50 million differential protein activities derived from drug perturbations *vs.* vehicle controls. These analyses revealed thousands of highly reproducible and novel, drug-mediated modulation of tissue-specific targets, leading to generation of a proteome-wide drug functional network, characterization of MoA-related drug clusters and off-target effects, dramatical expansion of druggable human proteome, and identification and experimental validation of novel, tissue-specific inhibitors of undruggable oncoproteins, most never reported before. The drug perturbation profile resource described here represents the first, large-scale, whole-genome-wide, RNA-Seq based dataset assembled to date, with the proposed framework, which is easily extended to elucidating the MoA of novel small-molecule libraries, facilitates mechanistic exploration of drug functions, supports systematic and quantitative approaches to precision oncology, and serves as a rich data foundation for drug discovery.

Introduction

Targeted therapy and precision medicine are predicated on the ability to accurately elucidate the mechanism of action (MoA) of small molecule compounds and to match it to disease-specific dependencies. However, while the binding affinity and molecular structure of targets revealed by standard biochemical assays (*on-target*) are typically well characterized, the efficacy and toxicity of drugs is being increasingly linked to their complex polypharmacology, as mediated by additional low-affinity binding proteins (*off-target*) and by tissue-specific indirect effector proteins (*secondary target*) that are poorly characterized or even undetectable by traditional assays[1]. Indeed, MoA characterization studies—including those based on recent advances in pull-down technology[2]—rely on *in vitro* affinity binding assays designed to investigate only specific protein classes, such as kinases[3], thus failing to reveal a compound's full range of high- and low-affinity binding targets and, perhaps more importantly, its tissue-specific effectors. Elucidating the proteome-wide mechanisms that determine drug's clinical efficacy and toxicity—as well as those that are pharmacologically irrelevant—has thus emerged as a critically-relevant, yet still elusive endeavor[3,4] in precision medicine, including in its oncology-related applications[1,5,6].

This rationale supported development of a framework aimed at addressing these challenges by integrating large-scale experimental assays and computational methodologies. The goal is to support a more rational approach to drug discovery or repurposing—including on an individual patient basis—by matching drug MoA to either pre-existing or *de novo* discovered tumor dependencies[7,8], across a wide range of aggressive[9,10] or orphan[11] human malignancies. Despite the oncology-specific nature of this study, the proposed framework is highly generalizable and can be used to study drug MoA in any physiologic or pathologic context (see Discussion).

Perturbational profiles, characterizing the response of a tissue to one or more small molecules, have been previously leveraged to study MoA. These include both proteomic assays[12-15], for

instance to elucidate signal-transduction protein targets[16], and transcriptional profiles focusing on changes in transcriptional cell state following drug perturbation[5,6,17]. Among these, proteomic assays are still expensive, labor intensive, and quite sparse. As such, they have been restricted to study only a handful of compounds of interest in specific cellular contexts, rather than large compound libraries[12,13], or to focus only on individual cell lines[14]. While transcriptional profiles are cheaper and more amenable to automation, their ability to directly inform on drug-mediated changes in protein activity or abundance is limited and their utilization has also been limited to small, representative gene sets[17] or to only a handful of tissues[6,18].

The methodology proposed in this manuscript leverages the extensively validated VIPER algorithm[19] to effectively transform large-scale, genome-wide transcriptomic profiles of drug perturbations into effective protein activity profiles. Akin to a highly multiplexed gene reporter assay, VIPER measures a protein's differential activity based on the differential expression of its tissue-specific targets (*regulon*) and has been shown to compare favorably to antibody-based measurements and to support an extensive set of monitored proteins (i.e., thousands vs. hundreds). The experimental component of this study is thus based on the high-throughput, low-cost generation of comprehensive, medium-depth (1 to 2 million reads) *transcriptional profile* libraries, using the PLATE-Seq technology[20], representing the response of multiple tissues to a comprehensive repertoire of clinically relevant and investigational compounds. The VIPER algorithm is then used to transform these profiles into proteome-wide *perturbational protein activity profiles,* comparing the effect of each drug vs. vehicle control on ~6,500 transcriptional regulators and signaling proteins.

While extensive perturbational profiles libraries, such as LINCS[17], have become recently available, they present limitations to effective elucidation of drug MoA, including their suitability to VIPER-based analyses. Specifically, (a) perturbations are generally performed at isomolar concentrations (e.g., 1μM, 5μM, etc.), thus potentially missing the relevant pharmacological

window where the compound's MoA is optimally manifested; (b) perturbed cell lines are not preselected based on their ability to faithfully represent the biology of human disease; (c) compound libraries have relatively low coverage of clinically relevant drugs, such as FDA-approved and late-stage developmental compounds (i.e., in phase 2 or 3 trials) and, critically, (d) these perturbational profiles cannot be used in VIPER-based analyses because their Luminex-based assays only quantify the expression of < 1,000 landmark genes, thus reducing protein-regulon representation by > 20-fold.

To address these challenges and to produce a large-scale repertoire of clinically-relevant perturbational profiles for their comprehensive polypharmacology characterization, we leveraged PLATE-Seq, a low-cost, fully-automated, genome-wide gene expression profiling methodology[20] yielding VIPER analyses that are virtually indistinguishable from those based on high-depth TruSeq profiles[19]. The resulting PAN-cancer Activity-based Compound Efficacy Analysis (PANACEA) dataset characterizes the proteome-wide effect of > 700 clinically relevant small molecules in an ever-growing number of cell lines (*n* = 23 in this manuscript). Critically, the cell lines used in this study were selected to represent high-fidelity models of specific human cancer subtypes, based on conservation of their most differentially active proteins, via the OncoMatch algorithm[9,21]. This is critical because we have shown that drug MoA predicted in such high-fidelity *in vitro* models is faithfully recapitulated in PDX models *in vivo*[22]. Finally, as shown in[6,23], to prevent activation of stress or cell death pathways that inevitably confound the MoA analysis and to avoid using concentrations that are too low to engage relevant MoA mechanisms, perturbational profiles were generated at each compound's highest sublethal concentration, as assessed from experimental drug response curves (DRC).

PANACEA profiles have been extremely effective in predicting drugs that abrogate tumor viability *in vivo*[9,22,24], suggesting they may be equally effective in supporting *de novo* drug characterization. This includes: (a) cellular network-mediated canalization of different drug perturbations, resulting

in activation of common response mechanisms, (b) proteome-wide assessment of both universal and tissue-specific MoA similarities, (c) elucidation of tissue-specific polypharmacology, and (d) discovery of novel, tissue-specific inhibitors of undruggable oncoproteins—including critical transcription factors and co-factors previously identified as Master Regulators of molecularly distinct tumor subtypes[25].

More specifically, in this manuscript we reported PANACEA identified drugs that could inhibit Master Regulator targets considered to be largely undruggable that were experimentally validated, such as CTTNNB1, UHRF1, MYC and CENPF. We also used PANACEA to predict compounds polypharmacology or assign MoA to compounds with yet-to-be-determined function, followed by experimental validations. For example, we found FLT3/ABL/Aurora kinase inhibitor KW2499 is also a PIK3CA/AKT pathway inhibitor. biib021 (an HSP90 inhibitor[26]) and az12419304 (an unknown MoA compound) can inhibit the activity of MEK pathway. Mgcd265, a multi-kinase inhibitor, also plays a role as DHFR pathway inhibitor.

The proposed approach is fully generalizable and can be adapted to studying the MoA of novel chemical libraries in any physiologic or pathologic context, from neurological and neurodegenerative diseases[27] to targeting mechanisms associated cell state reprograming[28] and viral infection[29]. The PANACEA resource is freely available to the research community.

Results

PANACEA Database Assembly: PANACEA comprises n = 18,046 RNASeq profiles representing the genome-wide transcriptional response of 23 cell lines at 24h following treatment with >700 FDA-approved and late-stage investigational compounds, as well as vehicle control media, in duplicate. The protocol to assemble this extensive dataset was designed with the objective of elucidating the tissue-specific mechanism of action (MoA) and polypharmacology of clinically relevant drugs to a level of detail currently unachievable by existing methodologies.

While PANACEA may appear conceptually similar to the Connectivity Map (CMap)[18] or LINCS (L1000)[17], there are fundamental differences critically required to achieve these goals. First, PANACEA comprises both genome-wide RNA-Seq profiles and protein activity profiles for ~6,500 regulatory and signaling proteins, as assessed by the extensively validated VIPER algorithm[19,22,24,30-32]. Second, PANACEA was specifically designed to support clinical applications by (a) selecting cell-lines representing high-fidelity models of rare and/or aggressive malignancy subtypes (**Figure 1A & Figure 1C**), and (b) including the vast majority of clinically-relevant oncology drugs—including FDA-approved and late-stage experimental agents in Phase 2 and 3 clinical trials (**Table S1**). Third, drug-concentration was chosen as the 48h, cell-line specific IC_{20}, as assessed from 10-point dose response curves (DRC) (**Figure 1B**). While this adds significant complexity and cost to the design, it is also critical to avoid confounding effects associated to stress/death pathway activation—accounting for up to 75% of the variance in existing drug-perturbation datasets—that confound MoA elucidation by masking critical targets, while artificially increasing drug MoA similarity[33,34].

The most critical advantage of this design, however, is that genome-wide perturbational profiles are optimally suited to network-based analyses (e.g., VIPER-based) aimed at reproducibly assessing drug-mediated effects on a complete repertoire of transcriptional regulators and signal

transduction proteins[19]. It should also be noted that, of the 23 PANACEA cell lines, only two (LoVo and BT20) are also included in LINCS[17]. This is important because PANACEA cell lines were selected to optimally represent the biology of human tumors based on their ability to represent high-fidelity models for TCGA (The Cancer Genome Atlas) cohorts, as well as additional cohorts representing rare and/or aggressive adult/pediatric malignancies.

Selecting patient-relevant cell line models for perturbational analysis: To support their ability to survive *ex vivo*, cell lines accumulate massive genetic, epigenetic, and aneuploidy artifacts that are absent in their parental tumors. While poorly characterized and understood, these events significantly affect their pathophysiologic fidelity, including their response to drugs. As a result, selecting cell lines for perturbational analysis based on the either (a) presence of specific driver mutations, (b) activity of specific surface markers, (c) *in vitro* growth kinetics, or (d) morphology may not represent the best strategy. For instance, the commonly used strategy of selecting cell lines with optimal growth kinetics can bias their sensitivity to cell cycle-dependent chemotoxic agents, while masking subtler dependencies that may be critical in slower-growing human tumors. Similarly, we have shown that virtually all but three of available neuroblastoma cell lines harboring *MYCN* amplifications fail to recapitulate the protein activity signature of *MYCN*-amplified tumors[31].

Several studies have shown that the Master Regulator (MR) proteins representing homeostatic determinants of tumor cell state—which can be effectively identified by VIPER analysis—are remarkably conserved within individual tumor subtypes, indeed much more so than mutational events[25]. Based on these observations, we have developed[9,31] and extensively validated[22,24,35] the OncoMatch algorithm, which assesses *model fidelity* based on the statistical significance of the MR overlap in a model/tumor pair. Specifically, OncoMatch tests for the enrichment the top 50 MRs—comprising the 25 most activated (MR ↑) and 25 most inactivated (MR ↓) proteins—in a patient tumor in proteins differentially active in a cancer model (e.g., cell line, PDX, organoid, etc.) and vice-versa. We have shown that this metric effectively recapitulates drug MoA—as assessed

in high-fidelity cell lines—in matched, low-passage PDXs[22]. Moreover, this metric is highly stable when the number of selected MRs varies between 10 and 200[22].

We thus used OncoMatch to optimize cell line selection for perturbational studies, with the objective of making them as relevant as possible to clinical translation. Specifically, we matched the gene expression profiles of thousands of cell lines from multiple collections—including the Cancer Cell Line Encyclopedia (CCLE)[36] and the Genentech Cell Line Screening Initiative (gCSI)[37], among others (see STAR Methods)—to assess their MR-based fidelity either to individual patients enrolled in our previously published N of 1 study at Columbia University[22] or to the largest fraction of samples in published, clinically relevant cancer cohorts (**Figure 1A**). For serous ovarian cancer, for instance, we prioritized high-fidelity models for two metastatic patients enrolled at Columbia University, leading to identification of the EFO21 cell line as the highest fidelity model (**Figure S8B**). This is important because the serous ovarian cancer cohort in TCGA does not include metastatic patients, whose MRs are significantly distinct from those identified in primary tumors. Similarly, for pediatric sarcoma, we use OncoMatch to assess the fidelity of available cell lines to 82 pediatric sarcoma patients samples collected at Memorial Sloan Kettering Cancer Center. The analysis identified SKNEP and MSKRMS as the two cell lines whose complementary matching produces the greatest coverage of that cohort (**Figure S9A**).

The 23 PANACEA cell lines include: H1793 (lung adenocarcinoma, LUAD), LNCAP and DU145 (AR-dependent and AR-independent prostate cancer, PRAD), PANC1 and ASPC1 (basal and classical pancreatic adenocarcinoma, PAAD), NLF (bladder cancer, BLCA), HCC1143 and BT20 (triple-negative breast cancer, TNBC), N87 (stomach, STAD), LOVO (colon cancer, COAD), HST-1 and KRJ1 (gastro-enteropancreatic neuroendocrine tumor, GEP-NET), IOMM (meningioma, ME), MSKOST and SAOS2 (pediatric osteosarcoma, OS), GIST (pediatric neuroblastoma, NBL), and HF2597 and U87 (proneural and mesenchymal glioblastoma, GBM), see STAR methods for details on individual cell line selection.

Perturbational profile generation: For each cell line, cells were first cultured and expanded using the recommended culture media (STAR Methods) until a sufficient number of cells was available to be transferred to 386-well plates for 10-point Dose Response Curve generation using a repertoire of 90 to 451 drugs (**Table S1**). Indeed, since the collection of these profiles has been ongoing for the last 10 years, the number and diversity of compounds per cell line is variable. However, a majority of these libraries includes ~140 FDA approved antineoplastics, ~170 late-stage experimental drugs, a variable number of tool compounds from the Tocris and MicroSource collections with $EC_{20} \leqslant 2\mu M$—as required for optimal utilization of 96- or 384-well PLATE-Seq plates—and vehicle control media (DMSO).

Cell lines were expanded until a sufficient number of cells was available for dispensing into 96- or 384-well plates, with an average of 4,000 cells per well (STAR methods). Cells were allowed to rest for 24h following by treatment with each compound at its cell line-specific 48h EC_{20}, as assessed by Dose Response Curves. Cells were then harvested at 24h following treatment and enzymatically processed for library preparation using the PLATE-Seq methodology to a depth of 1M to 2M reads per well. RNA-seq profiles were then processed for quality control and normalized (STAR methods).

Data Reproducibility: To assess gene expression profile reproducibility, we leveraged the LoVo and BT20 cell lines, which were included in both PANACEA and LINCS. As shown in **Figure S1 A-D**, the differences in Pearson and Spearman correlation between technical replicates (i.e., same cell line and drug), as assessed for the same set of ~1,000 LINCS landmark genes, were statistically significantly higher in PANACEA compared to LINCS, as assessed by the conservative, non-parametric Wilcoxon test. For LoVo, Wilcoxon p-values were $p = 2.75\times10^{-111}$ (Pearson) and $p = 6.77\times10^{-47}$ (Spearman), while, for BT20, they were $p = 4.71\times10^{-186}$ (Pearson) and $p = 4.39\times10^{-128}$ (Spearman). More critically, PANACEA includes whole-genome expression

profiles, while LINCS measures only ~1,000 landmark genes and uses computational methods to infer the expression of all other genes. The latter may be problematic since predictions are based on the analysis of cellular populations at steady state rather than following drug perturbation.

Next, as suggested by recent publications[25], we asked whether VIPER analyses could further improve reproducibility. Unfortunately, this analysis could not be performed for LINCS since use of 1,000 landmark genes decreases regulon representation by 20-fold. To use VIPER, the targets of regulatory (TFs/co-TFs) and signaling proteins were reversed engineered by analyzing RNA-seq profiles from histology-matched sample cohorts using the ARACNe algorithm[38] (see Table S2 for the corresponding tumor cohort list). First, for each cell line/drug pair, we measured the enrichment of genes that were significantly differentially expressed in one replicate (FDR < 0.1, Benjamini-Hochberg corrected) in genes that were differentially expressed or active in another replicate by Gene Set Enrichment Analysis (GSEA)[39] compared to vehicle control-treated samples. Then we performed the same analysis using VIPER-inferred protein activities. Specifically, we assessed the enrichment of the 50 most active and 50 most inactive proteins in drug vs. vehicle control-treated replicates. The results of this analysis are summarized for each of the 23 cell lines as violin plots representing the distribution across all drug perturbations (**Figure S2 A-W**). As assessed by Wilcoxon analysis, protein activity was significantly more reproducible than gene expression in all but one cell line, ranging from $p = 1.13\times10^{-13}$ in BT20 cells to $p = 1.53\times10^{-209}$ in SAOS2 cells. IOMM, a meningioma cell line, represented the only exception ($p = 0.24$) (**Figure S2 X**); this is not entirely surprising because, due to lack of large-scale, high-quality meningioma cohorts in literature; the interactome was generated using profiles from 106 samples collected at Columbia (Table S2), a sample size above the minimum required for ARACNe analysis and much smaller than for other represented malignancies.

Consistent with the fact that several drugs may be virtually inert in a specific cellular context—e.g., due to baseline inactivation of their high-affinity target—several cell lines presented violin

plots with a clear bi-modal structure. Indeed, their replicates will produce differential activity profiles that are essentially uncorrelated, resulting in a Normalized Enrichment Score (NES) distribution with median close to zero. In contrast, perturbations by more bioactive drugs produce higher NES values. Indeed, strong correlation of VIPER-assessed bioactivity and reproducibility was already shown in[9].

Drug Similarity Assessment: Consistent with these results—and based on prior studies showing the relevance of protein activity analysis in predicting drug sensitivity[9,25,32]—we leveraged protein activity profiles to elucidate drug MoA and polypharmacology.

To assess the MoA similarity between two drugs, Rx_A and Rx_B, we introduce a Drug Similarity Score ($DSS_{A\rightarrow B}$), representing the enrichment of the most differentially active proteins following treatment with Rx_A in protein differentially active following treatment with Rx_B, vs. vehicle control. The symmetric Drug Similarity Score metric ($DSS_{A,B}$) is then the average of $DSS_{A\rightarrow B}$ and $DSS_{B\rightarrow A}$ (**STAR methods**).

Critically, since drug MoA is tissue dependent, we also define a cell line MoA similarity score ($CMSS_{C_1\rightarrow C_2}$)—for the effect of the same drug in two cell lines (C_1 and C_2)—as the enrichment of the most differentially active proteins following C_1 perturbation in proteins differentially active following C_2 perturbation vs. vehicle control. The symmetric metric Cell Line MoA Similarity Score ($CMSS_{C_1,C_2}$) is then defined as the average of $CMSS_{C_1\rightarrow C_2}$ and $CMSS_{C_2\rightarrow C_1}$ (**STAR methods**).

Pleiotropy correction: Straightforward use of the DSS and CMSS metrics is plagued by poor specificity, because virtually unrelated drugs still induce pleiotropic activation of similar differential protein activity signatures. Investigation of this issue revealed that, while using a subtoxic concentration mitigates the issue, unrelated drugs still engage common cell-response mechanisms due to activation of common stress-response pathways. This contributes to a drug-

independent background signature of activated and inactivated (i.e., pleiotropic) proteins that translates into loss of MoA specificity. To remove such bias, we assigned a cell line-specific weight $w_{i,j} \in [0,1]$—where the i index identifies a specific protein and the j index a specific cell line—representing the overall conservation of its differential activity across all perturbational profiles for that cell line. The weight is computed using a sigmoidal transformation of the absolute NES value, for a specific protein and cell line, across all drugs treatments. This approach effectively mitigated most of the bias thus significantly improving MoA analysis specificity (STAR methods).

Cellular Context Specificity: Previous perturbational profile-based studies either assess compound MoA in a single cell line or integrate it across all cell lines perturbed with that specific compound[40]. As previously suggested, however, ignoring the tissue-specific nature of compound MoA may dramatically confounds the analysis[41]. Indeed, PANACEA profile analysis shows that individual drugs present highly conserved yet virtually orthogonal MoA in different cell line subsets, including in cell lines representing the same tumor subtype. For instance, as expected, resminostat and panobinostat—a class I/IIb and pan histone deacetylase (HDAC) inhibitor, respectively—present highly significant MoA similarity in the BT20 and HCC1143 TNBC cell lines ($p = 1.6\times10^{-9}$ and $p = 1.1\times10^{-12}$, respectively) (**Figure 2A & Figure 2B**). In contrast, pictilisib (a class I PI3K inhibitor) and neratinib (an ErbB family inhibitor) have highly significant conserved MoA in BT20 cells ($p = 9.6\times10^{-15}$) but not in HCC1143 cells ($p = 5.3\times10^{-1}$) (**Figure 2C & Figure 2D**), suggesting their MoA may be strongly cell-context dependent.

To further explore the tissue-specific nature of drug MoA, we assessed the enrichment of drug-pairs with the highest DSS score (top 1%) in one cell line in drug-pairs ranked by their DSS in a second cell line. Surprisingly, comparison of histologically distinct cell lines, such as U87 (GBM) and H1793 (LUAD), showed strikingly significant MoA conservation across a wide range of drugs ($p = 3.5\times10^{-12}$) (**Figure 2E**). In sharp contrast, histologically related cell lines, such as HCC1143 and BT20, both TNBC-derived, revealed much lower enrichment ($p = 9.6\times10^{-7}$) (**Figure 2F**).

MoA analysis must thus directly account for tissue-specificity, independent of cell histology. Indeed, averaging MoA across all available cell lines (including histology-matched ones) may dilute context-specific MoA similarities.

While this is consistent with expectations—since different cell lines may host different mutations, potentially bypassing or abrogating drug activity, or express different gene programs, including harboring or missing critical primary or secondary targets of the drug. Yet, a comprehensive map of such cell specific differences is still missing. For instance, the estrogen receptor (ESR1), an important pharmacological target in luminal breast cancer, has low or no expression in several PANACEA cell lines. As such, conservation of ESR1 inhibitors' MoA in these cells may derive from highly conserved off-target effects rather than their primary target. Unfortunately, since drug targets—including both high-affinity and off targets—are only partially characterized, the *a priori* selection of cell lines in which the MoA of a specific drug should be conserved is virtually impossible.

To address this challenge in unbiased, data driven fashion, we performed cluster analysis of each drug's MoA as assessed in each of the 23 available cell lines, thus revealing cell line clusters presenting orthogonal inter-cluster yet high intra-cluster MoA conservation patterns (STAR methods)—see, for instance, clustering results for crizotinib, imatinib, nilotinib, tamoxifen, topotecan and vandetanib (**Figure S3; Table S27).** As shown, the MoA of these drugs is highly conserved in different cell line subsets (positive DSS), with virtually opposite MoA between them (negative DSS), thus confirming an orthogonal subset of differentially active proteins, comprising primary and secondary targets, as well as effectors.

MoA Similarity Network: Even though some of the high-affinity binding targets of PANACEA drugs may be found in existing database (e.g., DrugBank[42]), the proteome-wide MoA and polypharmacology of these drugs are still largely unknown. In particular, DrugBank does not

include potential high-affinity binding targets and off-targets that are not assessed by traditional in vitro assays and ignores tissue specific secondary effectors.

To systematically elucidate the tissue-specific MoA conservation of tested drugs, we leveraged the previous section's analysis to identify MoA similarities as manifested at the level of each cell line cluster. Specifically, we leveraged hierarchical clustering to identify cell line subsets (clusters) with statistically significant intra-cluster drug MoA conservation, as assessed by the CMSS metric (**Figure S3**). Since clusters with statistically significant MoA conservation for two different drugs (Rx_A and Rx_B) may have variable overlap, we only considered cluster pairs with $\geqslant$ 50% overlap (i.e., Jaccard index $\geqslant$ 1/3) and with $\geqslant$ 2 cell lines each (STAR methods).

Each cluster intersection where the MoA of Rx_A and Rx_B was significantly conserved was then represented as an edge in a global MoA similarity network, with an associated p-value generated by integrating the DSS p-value of each cell line in the intersection, using Stouffer's method (see STAR methods for a relevant example). Such a network is difficult to effectively visualize and is better represented in tabular form (**Table S28;** STAR methods). For visualization purposes, we generated a simplified network with a single edge between drugs with statistically significant MoA conservation in one or more cell line subsets, associated with the p-value of the most significant edge in the complete network.

MoA Similarity Network Modularity Analysis: Network modularity analysis—such as used to identify protein complexes in protein-protein interaction network[43-48], for instance—represents an effective mean to identify subset of drugs (modules) presenting high intra-module and low-extra-module MoA conservation. In addition, this can be used to identify individual drugs with strong MoA conservation spanning across two or more modules, thus indicating potential polypharmacology. For this purpose, we leveraged the ClusterONE algorithm[49], followed by a

conservative statistical significance analysis (STAR methods). The analysis identified 30 distinct drug modules, associated with equally distinct pharmacologic properties (**Figure 3A & Table S3**).

To assess intra-module MoA conservation, we computed a Drug Modularity Score (DMS) p-value for each module drug as follows: first, we integrated the DSS-based p-values of the drug's MoA conservation with every other drug in the module, using Stouffer's method. The DMS representing the overall confidence that a drug belongs to a module. Then, to assess the statistical significance of the entire module, Module MoA Conservation Score ($MMCS$), we integrated the DMS of each module drug, also using Stouffer's method (STAR methods for null hypothesis models).

Based on literature analysis, a majority of the drug modules identified by the analysis were highly enriched in drugs with established MoA conservation. However, many drugs were identified with highly significant DMS, whose MoA was not previously reported as consistent with the module. This suggests that these drugs either present previously unsuspected polypharmacology or that tissue-specific cross-talk or canalization mechanisms induce pharmacological effects that are largely indistinguishable from those of other unrelated drugs, at least in a subset of the cell lines included in the study. Some key examples are shown in **Figure 3**. For example:

In Module M1, 13 of 14 drugs represent PI3K/AKT/mTOR axis inhibitors in Drug Bank (**Figure 3B)**, while the 14th drug, KW2499, represent a multi-kinase inhibitor, whose high-affinity binding targets include FLT3, ABL, ABL-T315I and Aurora kinase[50]. MoA analysis shows that KW2499 has highly significant DSS—in the range [10^{-27}, 10^{-326}], with the other drugs in the module, critically across most cell lines (i.e., 13 of 23 in the study). This cannot be discounted as an indirect effect of its primary targets because several other FLT/ABL/AURKA inhibitors are present in the database but are not included in the module (i.e., ENMD-2076). To experimentally validate the predicted polypharmacology of KW2499 as a PI3K/AKT/mTOR inhibitor, we measured the ratio of phosphorylated p70 ribosomal S6 kinase (p-P70S6K) by total p70 ribosomal S6 kinase (t-

P70S6K) after drug treatment in SAOS2 cells. It is noticed by the western blot that the level of p-70S6K was significantly reduced in the cells treated by KW2499 compared with the positive control 20% FBS, and the results from KW2499 treatment is consistent with Temsirolimus (a known mTOR inhibitor) (**Figure 4A-C**).

In **Module M2**, 13 of 14 drugs are established HDAC inhibitors (**Figure 3C**), while the 14th drug, cladribine, is an established adenosine analog that prevents the adenosine deaminase enzyme from modulating DNA processes[51]. Remarkably, its DSS with Belinostat and Rocilinostat—both established HDAC inhibitors—is even more significant ($p \leq 1.22\times10^{-136}$ and $p \leq 1.2\times10^{-105}$, respectively), than the DSS between Belinostat and Rocilinostat ($p \leq 5.21\times10^{-99}$), as supported by in cell lines PANC1, HSTS and GIST430. While the MoA of cladribine may suggest potential functional similarity to HDAC inhibitor-mediated chromatin remodeling, other adenosine and, more generally, purine analogs in PANACEA—e.g., pentostatin, clofarabine, mercaptopurine, and thioguanine, for instance—presented no statistically significant DSS with HDAC inhibitors, suggesting a potentially unique moonlighting function for this drug.

In **Module M3**, 17 of 22 drugs represent MAPK pathway inhibitors, including inhibitors of MEK and of its upstream modulators and downstream effector kinases (**Figure 3D**). The other 5 drugs represent a quite heterogeneous mixture, including az12419304 and az12456623—two small molecule inhibitors with a yet-to-be-determined function, which may thus represent novel MAPK inhibitors—biib021 (an HSP90 inhibitor[26]), danusertib (an AURKA inhibitor[52]), and ly2835219 (a CDK4/6 inhibitor[53]). To validate the prediction of this module, we used biib021 and az12419304 as examples. We measured the ratio of phosphorylated Erk (p-Erk) by total Erk (t-Erk) in N87 cells using western blot, and we observed the ratio was greatly reduced under treatment of biib021 and az12419304 compared with DMSO, which proved that biib021 and az12419304 work as MAPK pathway inhibitors (**Figure 4D-F**).

In **Module M18**, 12 of 13 drugs represent DNA cross-linking/alkylation and topoisomerase inhibitors (**Figure E**). However, the polo kinase (PLK1) inhibitor hmn214[54] was also predicted as a highly statistically significant member of M3 (ranging from $p \leq 5.26\times10^{-42}$ with irinotecan to 1.49×10^{-107} with fluorouracil), potentially consistent with its role as a modulator of DNA damage repair mechanisms.

Module M22 (**Figure F**) includes four established VEGF pathway inhibitors (sunitinib, axitinib, sorafenib, and cabozantinib), as well as two drugs not previously annotated as such, including amonafide—a topoisomerase inhibitor[55] that, surprisingly, failed to cluster with other topoisomerase inhibitors in M22—and bicalutamide, a selective androgen receptor blocker[56]. Both had highly significant DSS with VEGF pathway inhibitors. Specifically, amonafide's MoA similarity with sorafenib and cabozantinib was highly significant ($p \leq 1.03\times10^{-34}$ and $p \leq 6.22\times10^{-55}$, respectively), while bicalutamide similarity with axitinib and sorafenib was even more significant ($p \leq 1.36\times10^{-72}$ and $p \leq 4.64\times10^{-92}$, respectively).

As a final illustrative example, of the five drugs in **Module M24**, three (methotrexate, pemetrexed, and pralatrexate) are established dihydrofolate receptor (DHFR) inhibitors (**Figure G**), while a fourth one, floxuridine, is a thymidylate synthetase inhibitor[57], an enzyme also playing a key role in folate biosynthesis. However, Mgcd265 (also known as glesatinib) is an established multi-target kinase inhibitor[58] with no known relation to folate biosynthesis, thus establishing a potential link between these functions. To validate this prediction, we treated LoVo cells using Mgcd265 at various concentrations (0.5 μM, 1 μM and 5 μM, respectively) for 24 hours, as a result, reduction of DHFR at the protein level was observed for all three concentrations (**Figure 4G-I**).

Critically, modularity analysis also highlighted MoA conservation that is remarkably tissue-specific. For example, Module M29 includes four DNA damage drugs (mercaptopurine, ifosfamide, streptozotocin, and procarbazine) and two steroidogenesis inhibitors (mitotane and

aminoglutethimide) (**Figure S10A)**. While the two cytotoxic drug ifosfamide and mercaptopurine showed highly similar MoA across most cell lines, aminoglutethimide, mitotane, ifosfamide, procarbazine and streptozotocin formed a highly statistically significant sub-module only in three sarcoma-derived cell lines, including SAOS2 and MSKRMS (osteosarcoma) and SKNEP (Ewing sarcoma), again providing a sarcoma specific link between steroidogenesis and DNA-mediated cytotoxicity. An additional compelling example is provided by Module M13 (**Figure S10B)**, where drugs have highly distinct established MoA. While individual, statistically significant drug-drug edges are only supported by a handful of tissues, a highly interconnected sub-module, comprising five drugs (Gmx1778, Bay-117085, Cp100356, Ym201636 and Paroxetine) is identically conserved in the two TNBC cell lines (HCC1143 and BT20) suggesting a mechanism of drug effect canalization that is specific to this context. These illustrative examples show that tissue-specificity is a critical issue in drug MoA elucidation and suggest potential interesting tissue-specific biology for future investigation.

Targeting Undruggable Targets: Transcription factors and co-factors have emerged as some of the most critical tumor-essential proteins in many studies. Unfortunately, most of them are considered to be undruggable for lack of specific binding pockets[59] and lack of structural information. Indeed, for many of these proteins, only the DNA-binding domain(s) have been structurally resolved. We thus reasoned that the experimental perturbational profiles in PANACEA may offer critical insight into drugs that may inhibit these undruggable proteins, potentially as downstream effectors of a complex pharmacologic field-effect resulting from the polypharmacology-based, tissue-specific inhibition of multiple upstream modulators.

For instance, systematic pancancer analysis of ~10,000 samples in The Cancer Genomic Atlas (TCGA) elucidated Master Regulator (MR) proteins representing ultra-conserved, mechanistic determinants of tumor cell state, as well as the regulatory bottleneck responsible for canalizing the effect of patient-specific mutational landscapes. Critically, MRs have been shown to be highly

enriched in essential or synthetic lethal proteins, thus constituting a novel class of candidate drug targets[25,28,31,60,61]. Unfortunately, the majority of these proteins are considered to be undruggable.

A recent, pancancer study[25] identified 37 proteins as MRs of ⩾ 15 cancer subtypes each, thus establishing them as the most recurrent aberrantly activated MRs in TCGA. We thus interrogated the PANACEA repository to assess whether any of them could be inhibited by one of the tested drugs, in at least some of the evaluated tissue contexts in which they were identified as MRs. For this purpose, we first identified cell lines where these 37 proteins had above-average activity—i.e., VIPER score > 0 in untreated PANACEA cell lines, compared to the centroid of the CCLE and Genentech cell lines repositories[62]. We then analyzed the PANACEA perturbational profiles to identify drugs inducing statistically significant MR activity decrease in drug vs. vehicle control-treated cells. For each cell line, the top 15 drugs producing the strongest MR activity inhibition are shown as an NES heatmap (**Figure 5A)**. The three most potent, tissue-specific drugs targeting a cell line specific MRs are then recapitulated for each of the 23 PANACEA cell lines (**Figure 5B & Figure S4**).

Conclusions are two-fold: First, the top inhibitors of each MR were highly tissue-specific. This is not surprising, since the effect is likely mediated by the tissue-specific signaling and regulatory logic of the cell. Second, available drugs are only effective in targeting a subset of the selected MRs (**Figure 5B & Figure S4**). For example, statistically significant inhibitors of the OVOL1 MR could only be predicted for 4 of 23 cell lines. This suggests that use of VIPER-based analysis of perturbational profiles assays could be effectively leveraged to develop tissue-specific inhibitors for critical MRs or for any other protein with clinical translation relevance. These analyses also provide specificity assessment as VIPER can be used to determine whether a drug affects the activity of only a small vs. a large number of additional proteins.

An additional question was whether specific drugs may be identified as inhibitors of the same MR across multiple tissue contexts, thus suggesting a more direct mechanism. This is shown as a histogram showing the number of cell lines where the top 15 drugs for each MR were identified as statistically significant (**Figure 5C & Figure S5**). A table summarizing the effect of each drug on each MR, on a cell line by cell line basis, is also provided (**Table S4-26)**.

Experimental Validation of Tissue-specific, MR-targeting Drugs: In addition to mechanisms related to binding a protein's active site or to allosteric binding, protein activity modulation may result from three primary mechanisms, including: (a) indirect downregulation of the encoding gene's expression, (b) activation of ubiquitin-dependent proteasomal degradation, and (c) subcellular compartment sequestration (e.g., cytosol, nuclear/cell membrane, etc.). Since no single assay can reveal these complementary effects, we validated predicted MR inhibitors using three complementary assays, including (a) qRT-PCR, (b) Western Blot, and (c) confocal immunohistochemistry, respectively.

Selection of Drug/MR/cell line triplets for validation was largely based on availability of key reagents, such as MR-specific antibodies[25], cell line growth culture conditions, and baseline expression of the target. Four candidate MRs and predicted inhibitors/cell contexts were selected for Western-blot validation, including: (a) *CTNNB1 inhibition by the DNA intercalator dactinomycin in KRJ1 cells (GEP-NET)* (VIPER score = -4.17) (**Figure 6A**). β-catenin is a key effector of WNT signaling, which plays an important role in tumorigenesis[63]; (b) *UHRF1 inhibition by the HDAC inhibitor pracinostat in HF2597 cells (GBM)* (VIPER score = -11.07) (**Figure 6B**). UHRF1 plays an important role in the maintenance of DNA methylation in mammalian cells[64] has been identified as a potential drug target as it is essential to the survival of multiple human malignancies[65,66]; (c) *MYC inhibition by the MEK1/2 inhibitor pimasertib in HCC1143 (TNBC)* (VIPER score = -6.2) (**Figure 6C**). MYC is one of the most frequently amplified or translocated proto-oncogenes, associated with tumorigenesis in multiple of human maligancies[67], as well as a highly elusive drug

target; and (d) *CENPF inhibition by amsacrine in HF2597 (GBM)* (VIPER score = -12.8) (**Figure 6D**). CENPF is a large coiled-coil protein involved in centromere maintenance during mitosis[68]. In previous work, we have shown that co-inhibition of CENPF and FOXM1 abrogates viability of malignant castration resistant prostate cancer[69]. Indeed, these two proteins are co-aberrantly activated in several of the most aggressive malignancies in TCGA[25].

While Western Blot assays did not reveal overall changes in protein abundance, four additional MR-inhibitors could be validated by confocal microscopy of immunofluorescently stained cells, with statistical significance assessed by CellProfiler analysis using the conservative Wilcoxon test[70]. These include: (a) *Nuclear delocalization of CENPF in U87 cells (GBM)*, following treatment with the topoisomerase II inhibitor amsacrine (VIPER score = -12.31) (**Figure 6E**). Nuclear vs. cytoplasm localization of CENPF was significantly higher in in DMSO vs. amsacrine-treated samples ($p = 1.9\times10^{-6}$); (b) *Nuclear delocalization of CENPF in U87 cells (GBM)*, following treatment with the proteasome inhibitor bortezomib (VIPER score = -11.44) (**Figure 6E**). Nuclear vs. cytoplasm localization of CENPF was significantly higher in in DMSO vs. bortezomib-treated samples ($p = 3.8\times10^{-45}$); (c) *MYC sequestration to the nuclear lamina in LoVo cells (COAD)*, following treatment with the MEK1/2 inhibitor pimasertib (VIPER score = -6.32), consistent with prior reports of MYC activity inhibition by phosphorylation[71] (**Figure 6F**). Nuclear vs. nuclear-ring MYC localization was significantly higher in DMSO vs. pimasertib-treated samples ($p = 5.3\times10^{-186}$); and (d) *Nuclear delocalization of MYC in LoVo cells (COAD)*, following treatment with the MEK1/2 inhibitor refametinib (VIPER score = -8.23) (**Figure 6F**). Nuclear vs. cytoplasm MYC localization was significantly higher in in DMSO vs. bortezomib-treated samples ($p = 2.8\times10^{-63}$).

Finally, four additional MR-inhibitors downregulated the mRNA of the associated gene in the associated cell line context, as measured by qRT-PCR analysis, using three biological replicates and three technical replicates (**Figure 6G**). These include (a) *Bortezomib-mediated UHRF1*

*downregulation in H1793 cells (LUAD) (*VIPER score = -10.65*)*, (b) *Bortezomib-mediated CENPF downregulation in HF2597 cells (GBM) (*VIPER score = -6.18*)*, (c) *Pracinostat-mediated UHRF1 downregulation by in H1793 cells (LUAD) (*VIPER score = -11.21), and (d) *Bortezomib-mediated UHRF1 downregulation in HF2597 cells (GBM) (*VIPER score = -8.14).

Discussion

This study reports on a 10-year-long effort to create PANACEA, a collection of gene expression and protein activity profiles for the network-based elucidation of tumor-relevant drug mechanism of action and polypharmacology. For this purpose, PANACEA represents the genome/proteome-wide response of 23 cell lines, representing high-fidelity models of molecularly distinct cancer subtypes, to clinically relevant drug perturbations. Its design offers three main advantages: First, by leveraging the PLATE-Seq technology, cell response could be characterized on a genome-wide scale, both inexpensively and without substantial bias. Second, cell lines were chosen to represent clinically relevant models of human malignancies, based on OncoMatch analysis. This represents a critical aspect of this study as we have previously shown that MoA predictions based on PANACEA cell lines were recapitulated in vivo—in PDX models for which they were predicted as high-fidelity models by OncoMatch—for 15 of 18 tested drugs[22]. Third, drugs were carefully titrated at their highest sub-lethal concentration (EC_{20}) as assessed in each cell line, thus mitigating activation of cell stress and death pathways that may confound MoA analysis and preventing use of drugs at non bioactive concentrations. Most critically, availability of genome-wide perturbational profiles supports using the extensively validated VIPER algorithm to generate activity profiles for ~6,500 regulatory and signaling proteins, representing much more effective reporters of drug MoA, including secondary effectors and low-affinity binding targets may explain drug polypharmacology and toxicity.

VIPER assessment of differential protein activity in drug-treated cells is critical because most drugs do not affect the abundance of target proteins or the expression of encoding gene but rather their enzymatic or regulatory activity—as shown for estrogen receptor inhibitors[6,19], for instance. A critical advantage of VIPER is its high reproducibility, which has led to NY and CA Dept. of Health approval[72] of two VIPER-based algorithms, OncoTreat[9] and OncoTarget[10] for clinical use. Indeed VIPER-based protein activity measurements compare favorably with antibody-based

approaches, including at single cell level, where gene dropout severely compromises individual gene readouts[32,73]. VIPER prediction accuracy was confirmed in this study, since virtually all the drugs identified as potential context-specific Master Regulator inhibitors were experimentally validated, including via mechanisms, such as cell compartment localization, that would be difficult to assess by traditional assays.

Overall, the combination of the largest collection of genome/proteome-wide drug perturbation profiles and of the associated analytical framework provides a valuable resource to facilitate the identification of drugs targeting specific proteins, mechanisms, and pathways. This investigational framework can be easily extended to the study of small molecule compounds and biologics in virtually any disease-related or physiological context. For instance, we recently shown how this framework could be used to identify drugs, such as gemcitabine, that successfully reprogrammed the tumor tropism of regulatory T cells inducing dramatic synergy with immune checkpoint inhibitors, as well as predicting that optimal concentration would be 1/10th of what is currently used clinically[28]. In the context of drug discovery, VIPER analysis of drug perturbation profiles provides a universal phenotypic assay that, once generated, can be retooled at will to report on the effect of drugs on a variety of biologically relevant protein activity signatures—especially when compared to the typical highly-specialized single protein reporter assay used in drug discovery—thus providing deep characterization of an agent's pharmacology and toxicity for approximately the same cost per well.

The value of perturbational profiles in individual high-fidelity cell lines has been demonstrated in several previous studies. For instance, we have shown that OncoTreat analysis of a small subset of PANACEA cell lines could predict drugs that induce 91% disease control rate in PDX from patients who had failed multiple lines of therapy[22]. Similarly, OncoTreat analysis of HST-S and DU145 cell lines, was instrumental in the identification of drugs inducing partial response in vivo for gastroenteropancreatic neuroendocrine tumors (GEP-NETs) and castration resistant

metastatic prostate cancer (mCRPC), respectively. In addition, analysis of DU145 perturbation profiles identified four drugs capable of activating Master Regulator protein modules (MR-blocks) that are inactivated in the most aggressive subtype of prostate cancer, thus inhibiting growth, immune evasion, and cell migration hallmarks[25]. Supporting PANACEA's ability to extend to disease other than cancer, it should be noted that analysis of its perturbational profiles, for lung and colon cancer cell lines, was instrumental in identifying 18 drugs predicted to invert the activity if Master Regulator regulating host cell response to SARS-CoV-2 infection, 80% of which were experimentally validated[74]. Despite these success stories, however, none of these individual studies would have supported the large-scale analysis of MoA context-specificity nor the creation of a large-scale, tissue-specific MoA similarity and polypharmacology map discussed in this study.

Taken together, the results of this study suggest that PANACEA may provide a highly flexible and valuable framework to help achieve the ultimate goal of precision medicine, that is identifying small molecules targeting critical dependencies of both human malignancies and other human diseases.

Figure Legends

Figure 1: Schematic view of PANACEA resource for the characterization of oncology drug MoAs. (A) OncoMatch algorithm helps select the most clinically relevant cell lines from CCLE database using patient cohorts. (B) Determine IC_{20} for each drug in each cell line from 10-point dose response curves. (C) The table summarizes the cell lines, tissue of origins and number of perturbed drugs in PANACEA. (D) 23 selected cancer cell lines are perturbed by >700 oncology drugs for 24 hours at IC_{20}. (E) The PLATE-Seq technique was used to measure whole-genome wide gene expression after drug perturbation. (F) ARACNe and VIPER algorithms were used to compute protein activities after drug perturbations. (G) Calculate drug-drug functional association scores after grouping cell lines with similar drug effects. (H) Segment drug functional network to define drug functional modules for the inference of polypharmacology and prediction of drug mechanism of actions.

Figure 2: Enrichment analysis shows drug effects are cellular contexts dependent. (A) Two-tailed GSEA using resminostat induced protein activity vector (top and bottom 50 mostly differentiated proteins) against the ranked entire panobinostat induced protein activity vector in cell line BT20. (B) Two-tailed GSEA using resminostat induced protein activity vector (top and bottom 50 mostly differentiated proteins) against the ranked entire panobinostat induced protein activity vector in cell line HCC1143. (C) Two-tailed GSEA using pictilisib (Gdc0941) induced protein activity vector (top and bottom 50 mostly differentiated proteins) against the ranked entire neratinib induced protein activity vector in cell line BT20. (D) Two-tailed GSEA using pictilisib (Gdc0941) induced protein activity vector (top and bottom 50 mostly differentiated proteins) against the ranked entire neratinib induced protein activity vector in cell line HCC1143. (E) Enrichment analysis of top 1% drug-drug associations in cell line U87 vs. all statistically significant ($p < 0.01$, Bonferroni corrected) drug-drug associations in cell line H1793. U87 and H1793 have different tissues of origin. U87 is a glioblastoma cancer cell line and H1793 is a lung cancer cell

line. (F) Enrichment analysis of top 1% drug-drug associations in cell line HCC1143 vs. all statistically significant ($p < 0.01$, Bonferroni corrected) drug-drug associations in cell line BT20. Both HCC1143 and BT20 are triple negative breast cancer cell lines.

Figure 3: PANACEA derived drug functional map with highlighted examples. (A) The overview of all defined drug functional modules. A node represents one drug, and same node color indicates the same drug module assignment, and the node circle color is defined by the z-score of the drug using the Drug Specificity Test (STAR Methods, Table S3). Gray lines between drugs mean the drug-drug associations are conserved in at least two cell lines in PANACEA. (B-G) Six highlighted drug functional modules. Drugs marked using the same colors have similar annotated MoAs. Drugs colored by red indicate novel MoAs or polypharmacology.

Figure 4: Experimental validation of novel drug polypharmacology of MoA. (A) mTOR pathway inhibitors module, where red color indicates the drug is selected for experimental validation. (B) KW2449 inhibited the phosphorylation of p70 ribosomal S6 kinase (P70S6K), a downstream substrate phosphorylated by mTor, in a dose-dependent manner. Saos2 cells were treated for 6 hours with indicated concentrations of KW2449, Temsirolimus (a known inhibitor of mTor) or with 20% FBS as a positive control for mTor activation. Cell pellets were collected and lysed, and even amounts of proteins were separated by SDS-PAGE and blotted for phospho-P70S6 (p-P70S6), total-p70S6 (t-P70S6) and GAPDH as loading control. (C) Bar graph represents the quantification of the blot in (B). (D) MEK pathway inhibitors module, where red color indicates the selected drugs for experimental validation. (E) biib021 and az12419304 inhibited the phosphorylation of Erk. N87 cells were treated using biib021 or az12419304 at indicated concentrations for 4 hours. Cell pellets were collected and lysed, and even amounts of proteins were separated by SDS-PAGE and blotted for phospho-Erk (p-Erk), total-Erk (t-Erk) and tubulin as loading control. (F) Bar graph represents the quantification of the blot in (E). (G) DHFR pathway inhibitors module, where red color indicates the drug is selected for experimental

validation. (H) Mgcd265 induced a reduction in DHFR protein level. LoVo cells were treated with Mgcd265 at indicated concentrations for 24 hours. Cell pellets were collected and lysed, and even amounts of proteins were separated by SDS-PAGE and blotted for DHFR and tubulin (loading control). (I) Bar graph represents the quantification of the blot in (H).

Figure 5: A global view of PANACEA derived inhibitors for 37 most recurrent MRs. (A) A protein activity heatmap for most potent drug drugs for these 37 MRs in the 23 cancer cell lines screened in PANACEA database. (B) The top three most potent (ranked by VIPER score and VIPER score cut-off is <= -3) drugs for selected MRs across 23 cancer cell lines. (C) Bar charts show the 15 most potent drugs for selected MRs (ranked by the number of cell lines, where the drug has a VIPER score <= -3). In (B) and (c), VIPER score <= -3 indicates statistically significant MR activity inhibition in that cell line.

Figure 6: Experimental validation of novel inhibitors for transcriptional factors derived using PANACEA database. (A) Western blot result shows dactinomycin inhibits the activity of CTNNB1 in cell line KRJ1. (B) Western blot result shows pracinostat inhibits the activity of UHRF1 in cell line HF2597. (C) Western blot result shows pimasertib inhibits the activity of MYC in cell line HCC1143. (D) Western blot result shows amsacrine inhibits the activity of CENPF in cell line HF2597. (E) Confocal imaging analysis shows the signal of CENPF is greatly reduced in nucleus after treatment of amsacrine and bortezomib, which confirms our prediction that amsacrine and bortezomib could inhibit the activity of CENPF in cell line U87. (F) Confocal imaging analysis shows the signal of MYC is greatly reduced in nucleus after treatment of refametinib, and the signal of MYC is moved to nuclear lamina from nucleoplasm after treatment of pimasertib, which confirms our prediction that refametinib and pimasertib could inhibit the activity of CENPF in cell line U87. (G). Bar plots show the qRT-PCR results for four additional MR/drug/cell line predictions, and all predictions were confirmed by three biological replicates. The error bar is constructed by ± one standard deviation. The four heat maps in (A)-(D) give the top 15 most potent drugs for that

MR in corresponding cell line, and the validated drug names are marked using red color. The detailed chemical structures of mentioned drugs are also shown. The scale bar in (E) and (F) stands for 10 μm.

Acknowledgements

We acknowledge the High-Throughput Screening and the Genome Sequencing and Analysis Facilities at the JP Sulzberger Columbia Genome Center, the Confocal and Specialized Microscopy Shared Resource (SCMSR) managed by the Herbert Irving Comprehensive Cancer Center at Columbia University and supported in part by the NIH/NCI Cancer Center Support grant P30 CA 13696. This work was also supported by the NCI Office of Cancer Target Discovery and Development (CTD2) award U01CA272610, U54CA274506, and the NIH Shared Instrumentation Grants S10OD032433, S10OD012351 and S10OD021764, all to A.C.

Conflict of Interest

Dr. Califano is founder, equity holder, and consultant of DarwinHealth Inc., a company that has licensed some of the algorithms used in this manuscript from Columbia University. Columbia University is also an equity holder in DarwinHealth Inc.

Open Access Statement

This content of this study is published under open access subject to licensing agreements. The authors allow researchers to distribute and use data from the original source provided it is distributed free of cost for non-commercial purposes. Any profit-making is strictly prohibited.

References

1 Lin, A. *et al.* Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Science translational medicine* **11**, doi:10.1126/scitranslmed.aaw8412 (2019).

2 Meissner, F., Geddes-McAlister, J., Mann, M. & Bantscheff, M. The emerging role of mass spectrometry-based proteomics in drug discovery. *Nat Rev Drug Discov* **21**, 637-654, doi:10.1038/s41573-022-00409-3 (2022).

3 Klaeger, S. *et al.* The target landscape of clinical kinase drugs. *Science* **358**, doi:10.1126/science.aan4368 (2017).

4 Anighoro, A., Bajorath, J. & Rastelli, G. Polypharmacology: challenges and opportunities in drug discovery. *J Med Chem* **57**, 7874-7887, doi:10.1021/jm5006463 (2014).

5 Iorio, F. *et al.* Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc Natl Acad Sci U S A* **107**, 14621-14626, doi:10.1073/pnas.1000138107 (2010).

6 Woo, J. H. *et al.* Elucidating Compound Mechanism of Action by Network Perturbation Analysis. *Cell* **162**, 441-451, doi:10.1016/j.cell.2015.05.056 (2015).

7 Tsherniak, A. *et al.* Defining a Cancer Dependency Map. *Cell* **170**, 564-576 e516, doi:10.1016/j.cell.2017.06.010 (2017).

8 Boehm, J. S. *et al.* Cancer research needs a better map. *Nature* **589**, 514-516, doi:10.1038/d41586-021-00182-0 (2021).

9 Alvarez, M. J. *et al.* A precision oncology approach to the pharmacological targeting of mechanistic dependencies in neuroendocrine tumors. *Nat Genet* **50**, 979-989, doi:10.1038/s41588-018-0138-4 (2018).

10 Zeleke, T. Z. *et al.* Network-based assessment of HDAC6 activity predicts preclinical and clinical responses to the HDAC6 inhibitor ricolinostat in breast cancer. *Nat Cancer* **4**, 257-275, doi:10.1038/s43018-022-00489-5 (2023).

11 Coutinho, D. F. *et al.* Validation of a non-oncogene encoded vulnerability to exportin 1 inhibition in pediatric renal tumors. *Med (N Y)* **3**, 774-791 e777, doi:10.1016/j.medj.2022.09.002 (2022).

12 Blattmann, P. *et al.* Systems Pharmacology Dissection of Cholesterol Regulation Reveals Determinants of Large Pharmacodynamic Variability between Cell Lines. *Cell Syst* **5**, 604-619 e607, doi:10.1016/j.cels.2017.11.002 (2017).

13 Ruprecht, B. *et al.* A mass spectrometry-based proteome map of drug action in lung cancer cell lines. *Nat Chem Biol* **16**, 1111-1119, doi:10.1038/s41589-020-0572-3 (2020).

14 Mitchell, D. C. *et al.* A proteome-wide atlas of drug mechanism of action. *Nat Biotechnol*, doi:10.1038/s41587-022-01539-0 (2023).

15 Zecha, J. *et al.* Decrypting drug actions and protein modifications by dose- and time-resolved proteomics. *Science* **380**, 93-101, doi:10.1126/science.ade3925 (2023).

16 Lee, M. J. *et al.* Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. *Cell* **149**, 780-794, doi:10.1016/j.cell.2012.03.031 (2012).

17 Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437-1452 e1417, doi:10.1016/j.cell.2017.10.049 (2017).

18 Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929-1935, doi:313/5795/1929 [pii] 10.1126/science.1132939 (2006).
19 Alvarez, M. J. *et al.* Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat Genet* **48**, 838-847, doi:10.1038/ng.3593 (2016).
20 Bush, E. C. *et al.* PLATE-Seq for genome-wide regulatory network analysis of high-throughput screens. *Nature communications* **8**, 105, doi:10.1038/s41467-017-00136-z (2017).
21 Alvarez, M. J. *et al.* Reply to 'H-STS, L-STS and KRJ-I are not authentic GEPNET cell lines'. *Nat Genet* **51**, 1427-1428, doi:10.1038/s41588-019-0509-5 (2019).
22 Mundi, P. S. *et al.* A Transcriptome-Based Precision Oncology Platform for Patient-Therapy Alignment in a Diverse Set of Treatment-Resistant Malignancies. *Cancer discovery* **13**, 1386-1407, doi:10.1158/2159-8290.CD-22-1020 (2023).
23 Bansal, M. *et al.* A community computational challenge to predict the activity of pairs of compounds. *Nat Biotechnol* **32**, 1213-1222, doi:10.1038/nbt.3052 (2014).
24 Vasciaveo, A. *et al.* OncoLoop: A Network-Based Precision Cancer Medicine Framework. *Cancer discovery* **13**, 386-409, doi:10.1158/2159-8290.CD-22-0342 (2023).
25 Paull, E. O. *et al.* A modular master regulator landscape controls cancer transcriptional identity. *Cell* **184**, 334-351 e320, doi:10.1016/j.cell.2020.11.045 (2021).
26 Lundgren, K. *et al.* BIIB021, an orally available, fully synthetic small-molecule inhibitor of the heat shock protein Hsp90. *Mol Cancer Ther* **8**, 921-929, doi:10.1158/1535-7163.MCT-08-0758 (2009).
27 Shelanski, M. *et al.* A systems approach to drug discovery in Alzheimer's disease. *Neurotherapeutics* **12**, 126-131, doi:10.1007/s13311-014-0335-5 (2015).
28 Obradovic, A. *et al.* Systematic elucidation and pharmacological targeting of tumor-infiltrating regulatory T cell master regulators. *Cancer Cell* **41**, 933-949 e911, doi:10.1016/j.ccell.2023.04.003 (2023).
29 Laise, P. *et al.* The Host Cell ViroCheckpoint: Identification and Pharmacologic Targeting of Novel Mechanistic Determinants of Coronavirus-Mediated Hijacked Cell States. *bioRxiv*, doi:10.1101/2020.05.12.091256 (2020).
30 Ding, H. *et al.* Quantitative assessment of protein activity in orphan tissues and single cells using the metaVIPER algorithm. *Nature communications* **9**, 1471, doi:10.1038/s41467-018-03843-3 (2018).
31 Rajbhandari, P. *et al.* Cross-Cohort Analysis Identifies a TEAD4-MYCN Positive Feedback Loop as the Core Regulatory Element of High-Risk Neuroblastoma. *Cancer discovery* **8**, 582-599, doi:10.1158/2159-8290.CD-16-0861 (2018).
32 Obradovic, A. *et al.* Single-cell protein activity analysis identifies recurrence-associated renal tumor macrophages. *Cell* **184**, 2988-3005 e2916, doi:10.1016/j.cell.2021.04.038 (2021).
33 Szalai, B. *et al.* Signatures of cell death and proliferation in perturbation transcriptomics data-from confounding factor to effective prediction. *Nucleic Acids Res* **47**, 10010-10026, doi:10.1093/nar/gkz805 (2019).

34 Lukacisin, M. & Bollenbach, T. Emergent Gene Expression Responses to Drug Combinations Predict Higher-Order Drug Interactions. *Cell Syst* **9**, 423-433 e423, doi:10.1016/j.cels.2019.10.004 (2019).
35 Alvarez, M. J. *et al.* Unbiased Assessment of H-STS cells as high-fidelity models for gastro-enteropancreatic neuroendocrine tumor drug mechanism of action analysis. *bioRxiv* **677435**, doi:doi: https://doi.org/10.1101/677435 (2019).
36 Ghandi, M. *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503-508, doi:10.1038/s41586-019-1186-3 (2019).
37 Haverty, P. M. *et al.* Reproducible pharmacogenomic profiling of cancer cell line panels. *Nature* **533**, 333-337, doi:10.1038/nature17987 (2016).
38 Basso, K. *et al.* Reverse engineering of regulatory networks in human B cells. *Nat Genet* **37**, 382-390, doi:10.1038/ng1532 (2005).
39 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).
40 Zhu, J. *et al.* Prediction of drug efficacy from transcriptional profiles with deep learning. *Nat Biotechnol* **39**, 1444-1452, doi:10.1038/s41587-021-00946-z (2021).
41 Smith, I. & Haibe-Kains, B. Similarity bias from averaging signatures from the Connectivity Map. *bioRxiv*, 2022.2001.2024.477615, doi:10.1101/2022.01.24.477615 (2022).
42 Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* **46**, D1074-D1082, doi:10.1093/nar/gkx1037 (2018).
43 Krogan, N. J. *et al.* Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440**, 637-643, doi:10.1038/nature04670 (2006).
44 Havugimana, P. C. *et al.* A census of human soluble protein complexes. *Cell* **150**, 1068-1081, doi:10.1016/j.cell.2012.08.011 (2012).
45 Wan, C. *et al.* Panorama of ancient metazoan macromolecular complexes. *Nature* **525**, 339-344, doi:10.1038/nature14877 (2015).
46 Hu, L. Z. *et al.* EPIC: software toolkit for elution profile-based inference of protein complexes. *Nat Methods* **16**, 737-742, doi:10.1038/s41592-019-0461-4 (2019).
47 Xu, C. *et al.* Global Landscape of Native Protein Complexes in Synechocystis sp. PCC 6803. *Genomics Proteomics Bioinformatics* **20**, 715-727, doi:10.1016/j.gpb.2020.06.020 (2022).
48 Pourhaghighi, R. *et al.* BraInMap Elucidates the Macromolecular Connectivity Landscape of Mammalian Brain. *Cell Syst* **11**, 208, doi:10.1016/j.cels.2020.08.006 (2020).
49 Nepusz, T., Yu, H. & Paccanaro, A. Detecting overlapping protein complexes in protein-protein interaction networks. *Nat Methods* **9**, 471-472, doi:10.1038/nmeth.1938 (2012).
50 Shiotsu, Y. *et al.* KW-2449, a novel multikinase inhibitor, suppresses the growth of leukemia cells with FLT3 mutations or T315I-mutated BCR/ABL translocation. *Blood* **114**, 1607-1617, doi:10.1182/blood-2009-01-199307 (2009).
51 Leist, T. P. & Weissert, R. Cladribine: mode of action and implications for treatment of multiple sclerosis. *Clin Neuropharmacol* **34**, 28-35, doi:10.1097/WNF.0b013e318204cd90 (2011).
52 Gontarewicz, A. *et al.* Simultaneous targeting of Aurora kinases and Bcr-Abl kinase by the small molecule inhibitor PHA-739358 is effective against imatinib-resistant BCR-ABL

mutations including T315I. *Blood* **111**, 4355-4364, doi:10.1182/blood-2007-09-113175 (2008).

53 Gelbert, L. M. *et al.* Preclinical characterization of the CDK4/6 inhibitor LY2835219: in-vivo cell cycle-dependent/independent anti-tumor activities alone/in combination with gemcitabine. *Invest New Drugs* **32**, 825-837, doi:10.1007/s10637-014-0120-7 (2014).

54 Takagi, M. *et al.* In vivo antitumor activity of a novel sulfonamide, HMN-214, against human tumor xenografts in mice and the spectrum of cytotoxicity of its active metabolite, HMN-176. *Invest New Drugs* **21**, 387-399, doi:10.1023/a:1026282716250 (2003).

55 Hsiang, Y. H., Jiang, J. B. & Liu, L. F. Topoisomerase II-mediated DNA cleavage by amonafide and its structural analogs. *Mol Pharmacol* **36**, 371-376 (1989).

56 Furr, B. J. & Tucker, H. The preclinical development of bicalutamide: pharmacodynamics and mechanism of action. *Urology* **47**, 13-25; discussion 29-32, doi:10.1016/s0090-4295(96)80003-3 (1996).

57 Ferguson, P. J. *et al.* Antisense down-regulation of thymidylate synthase to suppress growth and enhance cytotoxicity of 5-FUdR, 5-FU and Tomudex in HeLa cells. *Br J Pharmacol* **127**, 1777-1786, doi:10.1038/sj.bjp.0702728 (1999).

58 Bradley, C. A. *et al.* Targeting c-MET in gastrointestinal tumours: rationale, opportunities and challenges. *Nat Rev Clin Oncol* **14**, 562-576, doi:10.1038/nrclinonc.2017.40 (2017).

59 Henley, M. J. & Koehler, A. N. Advances in targeting 'undruggable' transcription factors with small molecules. *Nat Rev Drug Discov* **20**, 669-688, doi:10.1038/s41573-021-00199-0 (2021).

60 Aytes, A. *et al.* Cross-species regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy. *Cancer Cell* **25**, 638-651, doi:10.1016/j.ccr.2014.03.017 (2014).

61 Carro, M. S. *et al.* The transcriptional network for mesenchymal transformation of brain tumours. *Nature* **463**, 318-325, doi:nature08712 [pii] 10.1038/nature08712 (2010).

62 Klijn, C. *et al.* A comprehensive transcriptional portrait of human cancer cell lines. *Nat Biotechnol* **33**, 306-312, doi:10.1038/nbt.3080 (2015).

63 Anastas, J. N. & Moon, R. T. WNT signalling pathways as therapeutic targets in cancer. *Nat Rev Cancer* **13**, 11-26, doi:10.1038/nrc3419 (2013).

64 Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760-1764, doi:10.1126/science.1147939 (2007).

65 Unoki, M., Brunet, J. & Mousli, M. Drug discovery targeting epigenetic codes: the great potential of UHRF1, which links DNA methylation and histone modifications, as a drug target in cancers and toxoplasmosis. *Biochem Pharmacol* **78**, 1279-1288, doi:10.1016/j.bcp.2009.05.035 (2009).

66 Kostyrko, K. *et al.* UHRF1 is a mediator of KRAS driven oncogenesis in lung adenocarcinoma. *Nature communications* **in press** (2023).

67 McKeown, M. R. & Bradner, J. E. Therapeutic strategies to inhibit MYC. *Cold Spring Harb Perspect Med* **4**, doi:10.1101/cshperspect.a014266 (2014).

68 Feng, J., Huang, H. & Yen, T. J. CENP-F is a novel microtubule-binding protein that is essential for kinetochore attachments and affects the duration of the mitotic checkpoint delay. *Chromosoma* **115**, 320-329, doi:10.1007/s00412-006-0049-5 (2006).

69 Aytes, A. *et al.* Cross-species regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy. *Cancer Cell* **25**, 638-651, doi:10.1016/j.ccr.2014.03.017 (2014).

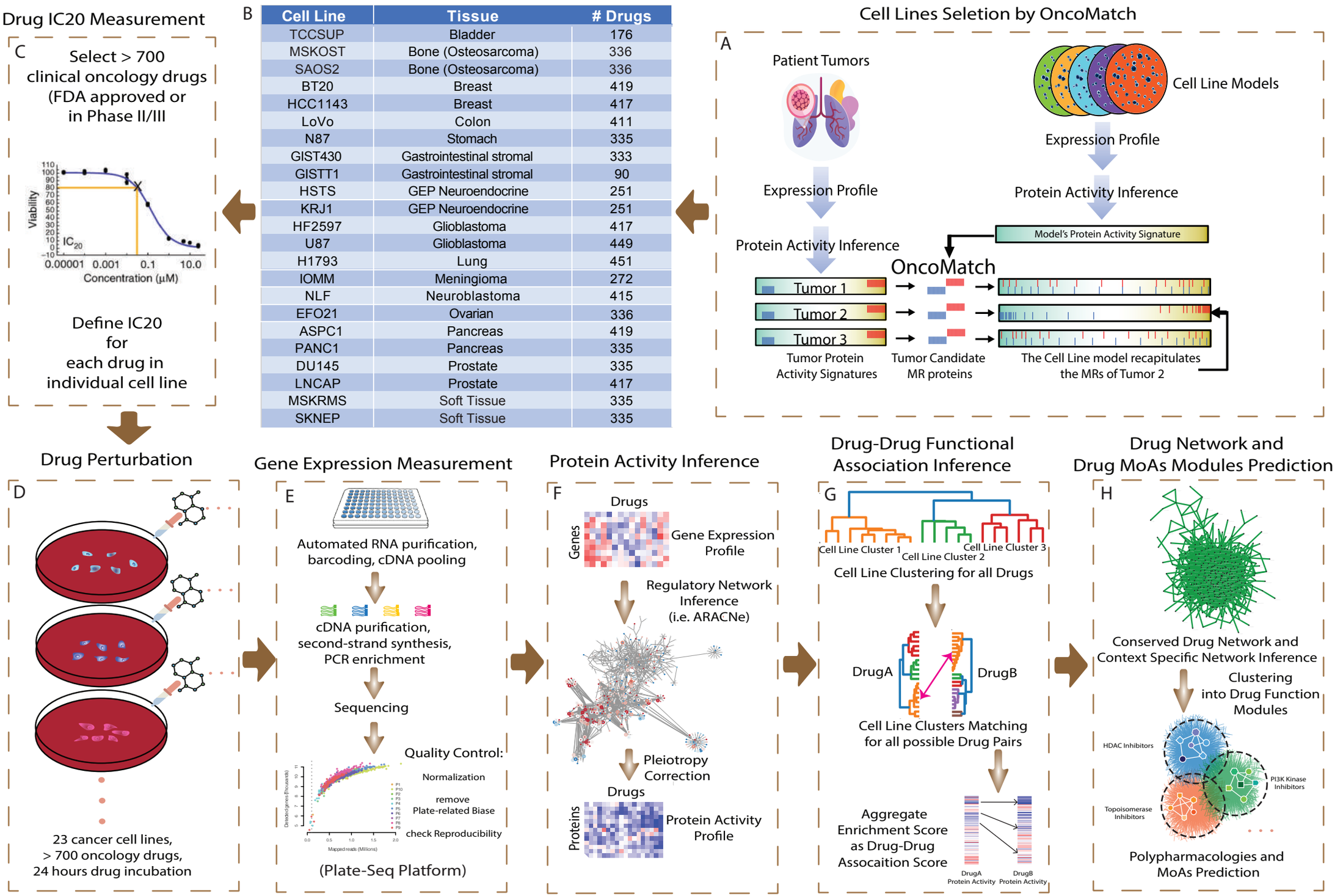
70 Stirling, D. R. *et al.* CellProfiler 4: improvements in speed, utility and usability. *Bmc Bioinformatics* **22**, 433, doi:10.1186/s12859-021-04344-9 (2021).

71 Myant, K. *et al.* Serine 62-Phosphorylated MYC Associates with Nuclear Lamins and Its Regulation by CIP2A Is Essential for Regenerative Proliferation. *Cell Rep* **12**, 1019-1031, doi:10.1016/j.celrep.2015.07.003 (2015).

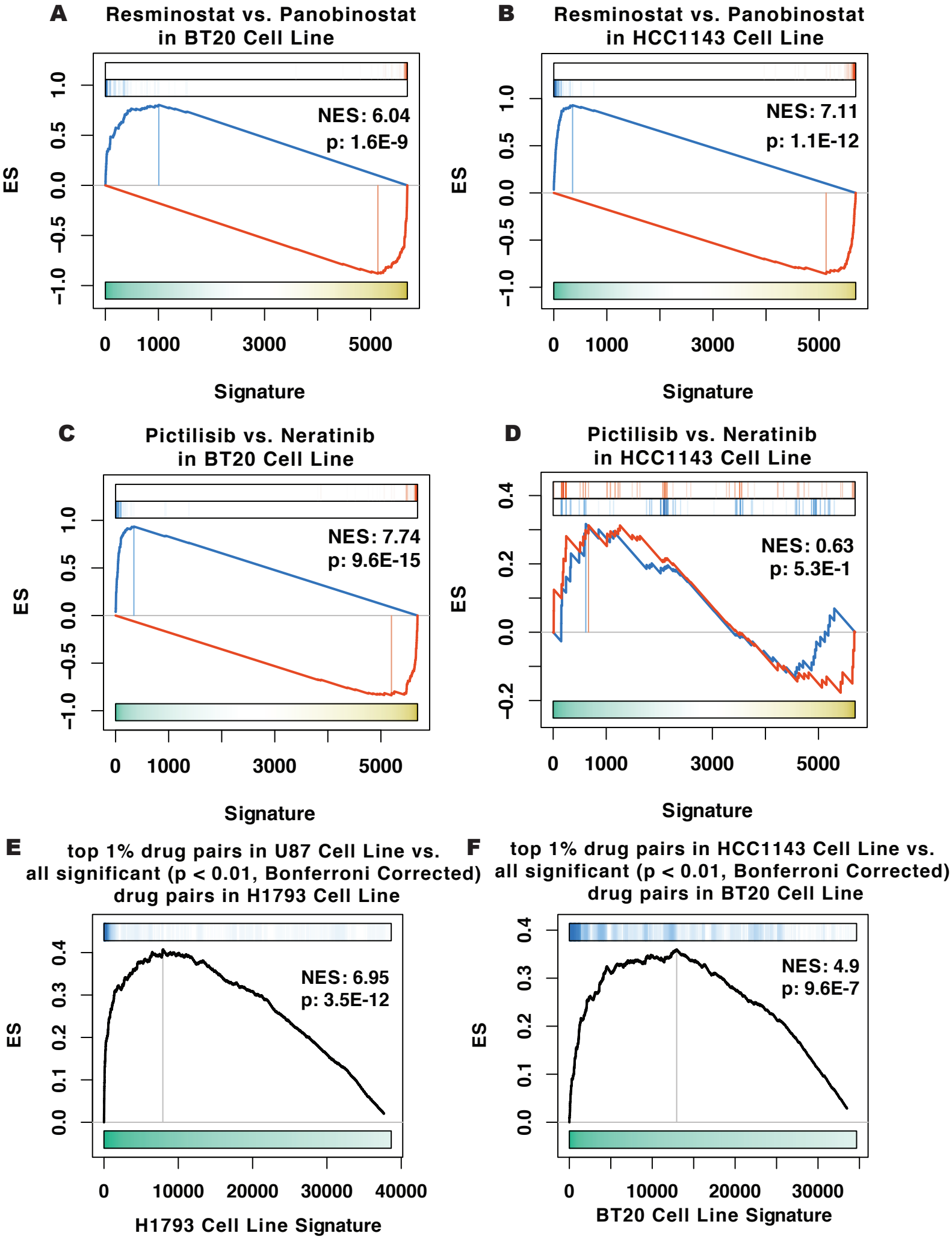
72 Alvarez, M. J. & Califano, A. *Darwin OncoTarget/OncoTreat: NY CLIA certified tests to identify effective drugs on an individual cancer patient basis from RNASeq profiles* <https://www.pathology.columbia.edu/diagnostic-specialties/personalized-genomic-medicine/oncology-testing/darwin-oncotarget-tm-oncotreat> (2018).

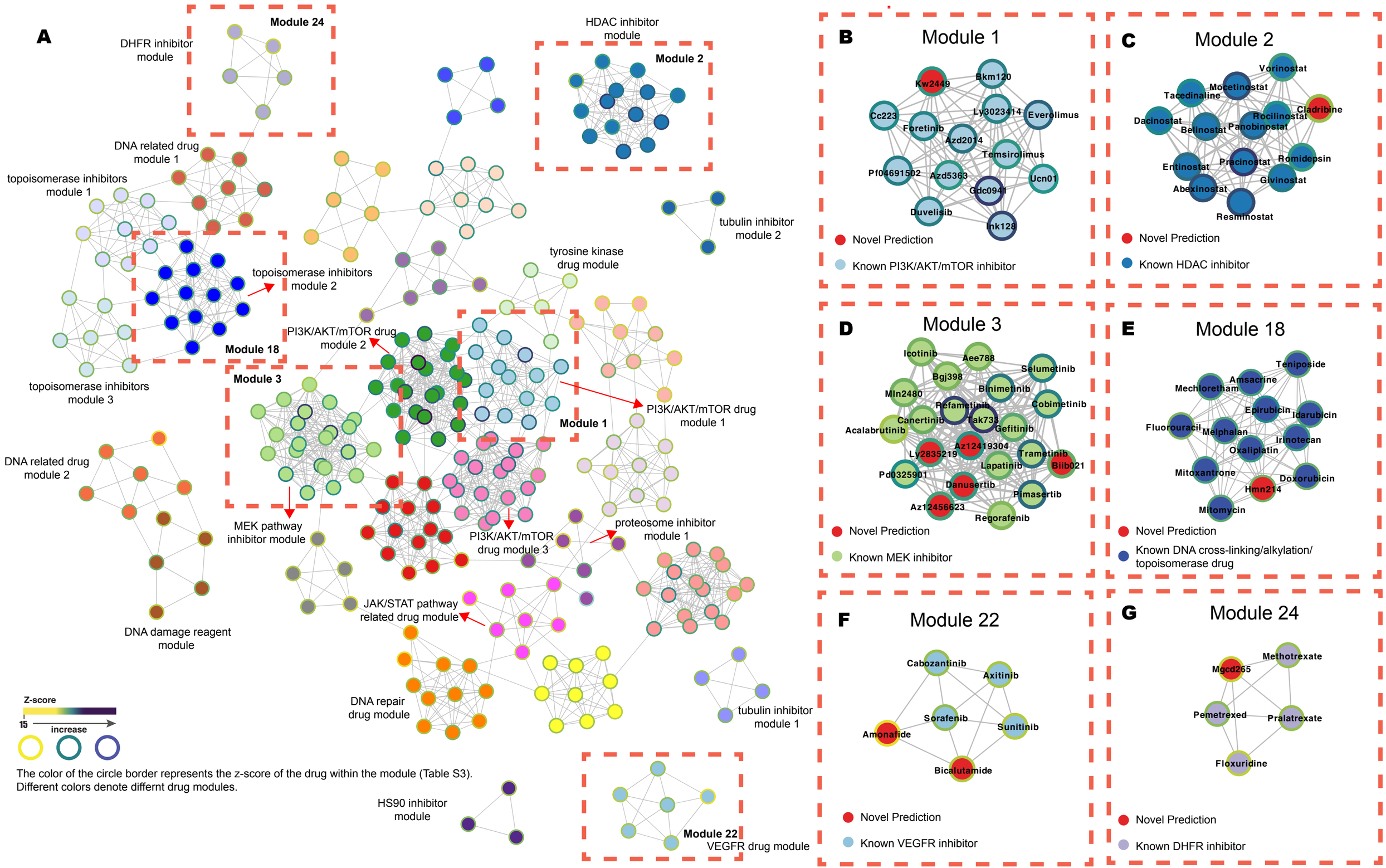
73 Vlahos, L. *et al.* Systematic, Protein Activity-based Characterization of Single Cell State. *bioRxiv* **bioRxiv 2021.05.20.445002**, doi:https://doi.org/10.1101/2021.05.20.445002 (2023).

74 Laise, P. *et al.* A model for network-based identification and pharmacological targeting of aberrant, replication-permissive transcriptional programs induced by viral infection. *Commun Biol* **5**, 714, doi:10.1038/s42003-022-03663-8 (2022).

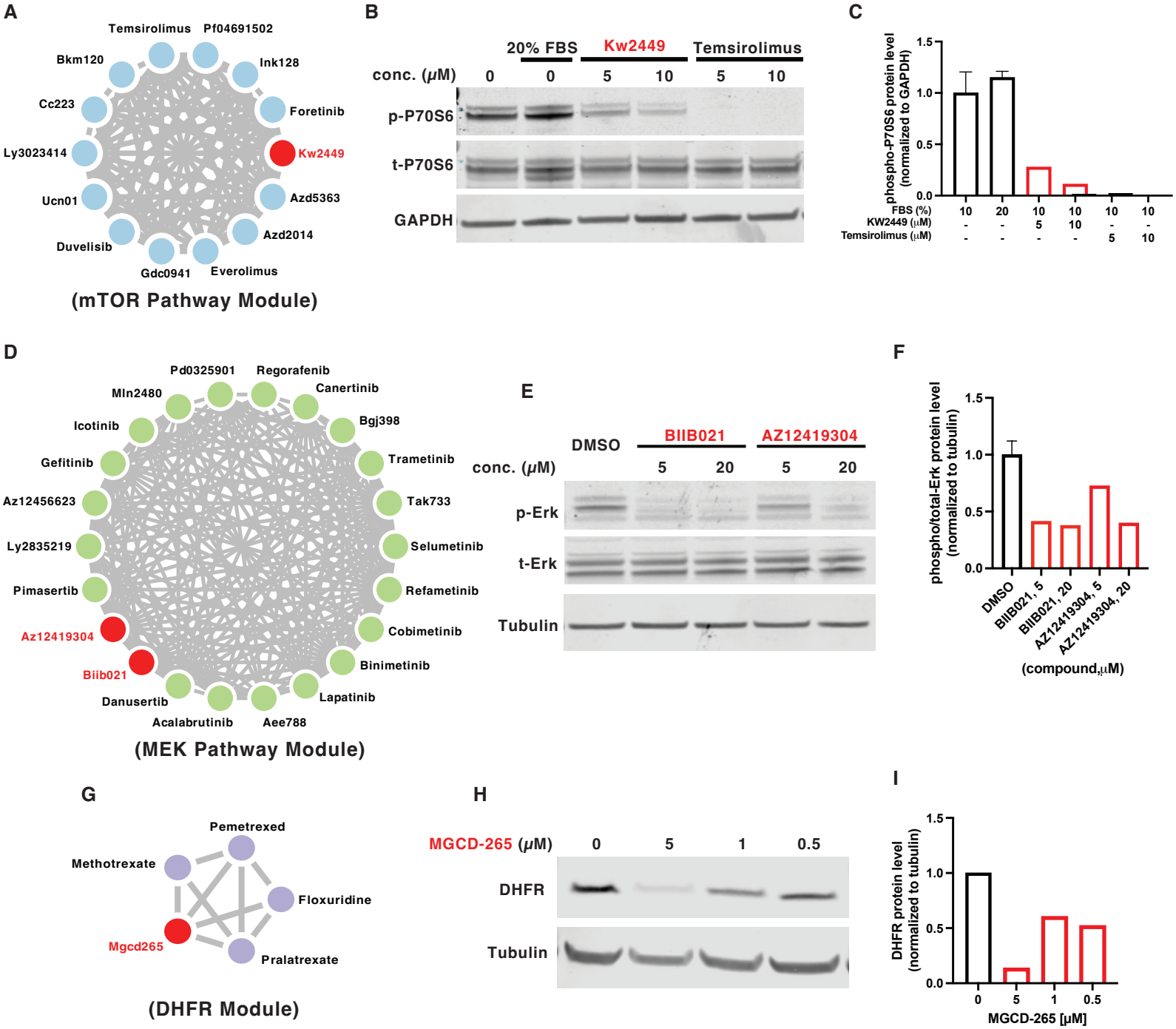


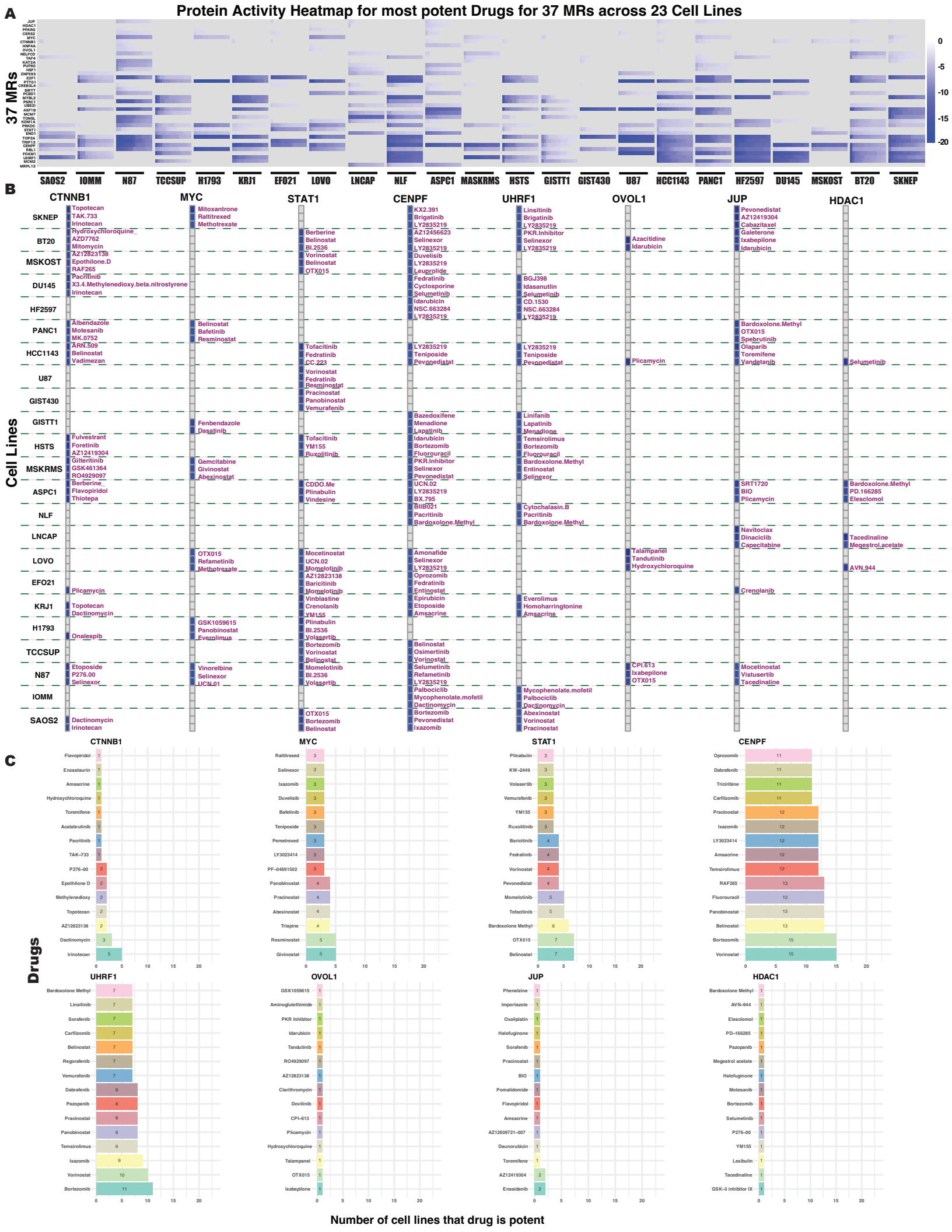
Cell Line	Tissue	# Drugs
TCCSUP	Bladder	176
MSKOST	Bone (Osteosarcoma)	336
SAOS2	Bone (Osteosarcoma)	336
BT20	Breast	419
HCC1143	Breast	417
LoVo	Colon	411
N87	Stomach	335
GIST430	Gastrointestinal stromal	333
GISTT1	Gastrointestinal stromal	90
HSTS	GEP Neuroendocrine	251
KRJ1	GEP Neuroendocrine	251
HF2597	Glioblastoma	417
U87	Glioblastoma	449
H1793	Lung	451
IOMM	Meningioma	272
NLF	Neuroblastoma	415
EFO21	Ovarian	336
ASPC1	Pancreas	419
PANC1	Pancreas	335
DU145	Prostate	335
LNCAP	Prostate	417
MSKRMS	Soft Tissue	335
SKNEP	Soft Tissue	335

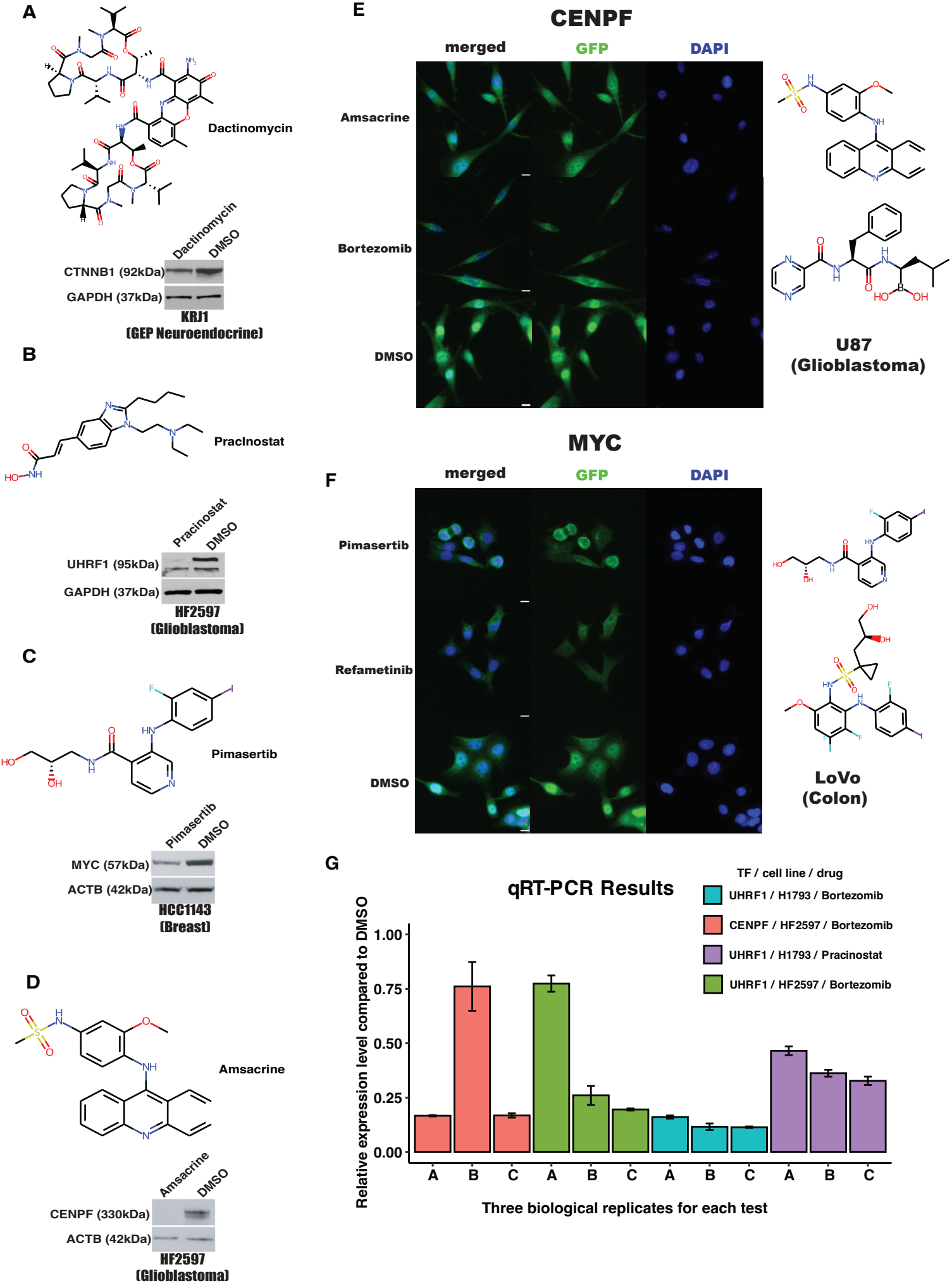

A
Resminostat vs. Panobinostat
in BT20 Cell Line
NES: 6.04
p: 1.6E-9
ES
Signature
B
Resminostat vs. Panobinostat
in HCC1143 Cell Line
NES: 7.11
p: 1.1E-12
ES
Signature
C
Pictilisib vs. Neratinib
in BT20 Cell Line
NES: 7.74
p: 9.6E-15
ES
Signature
D
Pictilisib vs. Neratinib
in HCC1143 Cell Line
NES: 0.63
p: 5.3E-1
ES
Signature
E
top 1% drug pairs in U87 Cell Line vs.
all significant (p < 0.01, Bonferroni Corrected)
drug pairs in H1793 Cell Line
NES: 6.95
p: 3.5E-12
ES
H1793 Cell Line Signature
F
top 1% drug pairs in HCC1143 Cell Line vs.
all significant (p < 0.01, Bonferroni Corrected)
drug pairs in BT20 Cell Line
NES: 4.9
p: 9.6E-7
ES
BT20 Cell Line Signature



The color of the circle border represents the z-score of the drug within the module (Table S3). Different colors denote differnt drug modules.


A
Temsirolimus
Pf04691502
Bkm120
Ink128
Cc223
Foretinib
Ly3023414
Kw2449
Ucn01
Azd5363
Duvelisib
Azd2014
Gdc0941
Everolimus
(mTOR Pathway Module)
B
20% FBS
Kw2449
Temsirolimus
conc. (μM)
0
0
5
10
5
10
p-P70S6
t-P70S6
GAPDH
C
phospho-P70S6 protein level (normalized to GAPDH)
1.5
1.0
0.5
0.0
FBS (%)
KW2449 (μM)
Temsirolimus (μM)
D
Pd0325901
Regorafenib
Mln2480
Canertinib
Icotinib
Bgj398
Gefitinib
Trametinib
Az12456623
Tak733
Ly2835219
Selumetinib
Pimasertib
Refametinib
Az12419304
Cobimetinib
Biib021
Binimetinib
Danusertib
Lapatinib
Acalabrutinib
Aee788
(MEK Pathway Module)
E
DMSO
BIIB021
AZ12419304
conc. (μM)
5
20
5
20
p-Erk
t-Erk
Tubulin
F
phospho/total-Erk protein level (normalized to tubulin)
1.5
1.0
0.5
0.0
DMSO
BIIB021, 5
BIIB021, 20
AZ12419304, 5
AZ12419304, 20
(compound,μM)
G
Pemetrexed
Methotrexate
Floxuridine
Mgcd265
Pralatrexate
(DHFR Module)
H
MGCD-265 (μM)
0
5
1
0.5
DHFR
Tubulin
I
DHFR protein level (normalized to tubulin)
1.5
1.0
0.5
0.0
0
5
1
0.5
MGCD-265 [μM]


Protein Activity Heatmap for most potent Drugs for 37 MRs across 23 Cell Lines
A
37 MRs
JUP
HDAC1
PPARG
CERS2
MYC
CTNNB1
HNF4A
OVOL1
NELFCD
TAF4
KAT2A
PUF60
HSF1
ZNF593
E2F1
PTTG1
CREB3L4
SRT7
RXRA
MYBL2
PSRC1
USF2I
ASF1B
MCM7
TONSL
KDM1A
PRKDC
STAT1
ENO1
TOP2A
TBP13
CENPF
RBL1
FOXM1
UHRF1
MCM2
MRPL12
0
-5
-10
-15
-20
SAOS2
IOMM
N87
TCCSUP
H1793
KRJ1
EFO21
LOVO
LNCAP
NLF
ASPC1
MASKRMS
HSTS
GISTT1
GIST430
U87
HCC1143
PANC1
HF2597
DU145
MSKOST
BT20
SKNEP
B
Cell Lines
CTNNB1
MYC
STAT1
CENPF
UHRF1
OVOL1
JUP
HDAC1
SKNEP
Topotecan
TAK.733
Irinotecan
Mitoxantrone
Raltitrexed
Methotrexate
BT20
Hydroxychloroquine
AZD7762
Mitomycin
Berberine
Belinostat
BI.2536
Vorinostat
Belinostat
OTX015
KX2.391
Brigatinib
LY2835219
AZ12456623
Selinexor
LY2835219
Duvelisib
LY2835219
Leuprolide
Fedratinib
Cyclosporine
Selumetinib
Idarubicin
NSC.663284
LY2835219
Linsitinib
Brigatinib
LY2835219
PKR.Inhibitor
Selinexor
LY2835219
Pevonedistat
AZ12419304
Cabazitaxel
Galeterone
Ixabepilone
Idarubicin
Azacitidine
Idarubicin
MSKOST
AZ12823138
Epothilone.D
RAF265
DU145
Pacritinib
X3.4.Methylenedioxy.beta.nitrostyrene
Irinotecan
BGJ398
Idasanutlin
Selumetinib
CD.1530
NSC.663284
LY2835219
HF2597
PANC1
Albendazole
Motesanib
MK.0752
Belinostat
Bafetinib
Resminostat
Bardoxolone.Methyl
OTX015
Spebrutinib
Olaparib
Toremifene
Vandetanib
HCC1143
ARN.509
Belinostat
Vadimezan
Tofacitinib
Fedratinib
CC.223
LY2835219
Teniposide
Pevonedistat
LY2835219
Teniposide
Pevonedistat
Plicamycin
Selumetinib
U87
Vorinostat
Fedratinib
Resminostat
Pracinostat
Panobinostat
Vemurafenib
GIST430
GISTT1
Fenbendazole
Dasatinib
Bazedoxifene
Menadione
Lapatinib
Idarubicin
Bortezomib
Fluorouracil
PKR.Inhibitor
Selinexor
Pevonedistat
Linifanib
Lapatinib
Menadione
Temsirolimus
Bortezomib
Fluorouracil
Bardoxolone.Methyl
Entinostat
Selinexor
HSTS
Fulvestrant
Foretinib
AZ12419304
Tofacitinib
YM155
Ruxolitinib
MSKRMS
Gilteritinib
GSK461364
RO4929097
Gemcitabine
Givinostat
Abexinostat
ASPC1
Berberine
Flavopiridol
Thiotepa
CDDO.Me
Plinabulin
Vindesine
UCN.02
LY2835219
BX.795
BIIB021
Pacritinib
Bardoxolone.Methyl
Cytochalasin.B
Pacritinib
Bardoxolone.Methyl
SRT1720
BIO
Plicamycin
Bardoxolone.Methyl
PD.166285
Elesclomol
NLF
LNCAP
Navitoclax
Dinaciclib
Capecitabine
Tacedinaline
Megestrol.acetate
LOVO
OTX015
Refametinib
Methotrexate
Mocetinostat
UCN.02
Momelotinib
Amonafide
Selinexor
LY2835219
Talampanel
Tandutinib
Hydroxychloroquine
AVN.944
EFO21
Plicamycin
AZ12823138
Baricitinib
Momelotinib
Oprozomib
Fedratinib
Entinostat
Crenolanib
KRJ1
Topotecan
Dactinomycin
Vinblastine
Crenolanib
YM155
Epirubicin
Etoposide
Amsacrine
Everolimus
Homoharringtonine
Amsacrine
H1793
GSK1059615
Panobinostat
Everolimus
Plinabulin
BI.2536
Onalespib
Volasertib
TCCSUP
Bortezomib
Vorinostat
Belinostat
Belinostat
Osimertinib
Vorinostat
Selumetinib
Refametinib
LY2835219
N87
Etoposide
P276.00
Selinexor
Vinorelbine
Selinexor
UCN.01
Momelotinib
BI.2536
Volasertib
CPI.613
Ixabepilone
OTX015
Mocetinostat
Vistusertib
Tacedinaline
IOMM
Palbociclib
Mycophenolate.mofetil
Dactinomycin
Mycophenolate.mofetil
Palbociclib
Dactinomycin
Abexinostat
Vorinostat
Pracinostat
SAOS2
Dactinomycin
Irinotecan
OTX015
Bortezomib
Belinostat
Bortezomib
Pevonedistat
Ixazomib
C
Drugs
CTNNB1
Flavopiridol 1
Enzastaurin 1
Amsacrine 1
Hydroxychloroquine 1
Toremifene 1
Acalabrutinib 1
Pacritinib 1
TAK−733 1
P276−00 2
Epothilone D 2
Methylenedioxy 2
Topotecan 2
AZ12823138 2
Dactinomycin 3
Irinotecan 5
MYC
Raltitrexed 3
Selinexor 3
Ixazomib 3
Duvelisib 3
Bafetinib 3
Teniposide 3
Pemetrexed 3
LY3023414 3
PF−04691502 3
Panobinostat 4
Pracinostat 4
Abexinostat 4
Triapine 4
Resminostat 5
Givinostat 5
STAT1
Plinabulin 3
KW−2449 3
Volasertib 3
Vemurafenib 3
YM155 3
Ruxolitinib 3
Baricitinib 4
Fedratinib 4
Vorinostat 4
Pevonedistat 4
Momelotinib 5
Tofacitinib 5
Bardoxolone Methyl 6
OTX015 7
Belinostat 7
CENPF
Oprozomib 11
Dabrafenib 11
Triciribine 11
Carfilzomib 11
Pracinostat 12
Ixazomib 12
LY3023414 12
Amsacrine 12
Temsirolimus 12
RAF265 13
Fluorouracil 13
Panobinostat 13
Belinostat 13
Bortezomib 15
Vorinostat 15
UHRF1
Bardoxolone Methyl 7
Linsitinib 7
Sorafenib 7
Carfilzomib 7
Belinostat 7
Regorafenib 7
Vemurafenib 7
Dabrafenib 8
Pazopanib 8
Pracinostat 8
Panobinostat 8
Temsirolimus 8
Ixazomib 9
Vorinostat 10
Bortezomib 11
OVOL1
GSK1059615 1
Aminoglutethimide 1
PKR Inhibitor 1
Idarubicin 1
Tandutinib 1
RO4929097 1
AZ12823138 1
Clarithromycin 1
Dovitinib 1
CPI−613 1
Plicamycin 1
Hydroxychloroquine 1
Talampanel 1
OTX015 1
Ixabepilone 1
JUP
Phenelzine 1
Importazole 1
Oxaliplatin 1
Halofuginone 1
Sorafenib 1
Pracinostat 1
BIO 1
Pomalidomide 1
Flavopiridol 1
Amsacrine 1
AZ12609721−007 1
Daunorubicin 1
Toremifene 1
AZ12419304 2
Enasidenib 2
HDAC1
Bardoxolone Methyl 1
AVN−944 1
Elesclomol 1
PD−166285 1
Pazopanib 1
Megestrol acetate 1
Halofuginone 1
Motesanib 1
Bortezomib 1
Selumetinib 1
P276−00 1
YM155 1
Lexibulin 1
Tacedinaline 1
GSK−3 inhibitor IX 1
0 5 10 15 20
Number of cell lines that drug is potent


A
Dactinomycin
Dactinomycin
DMSO
CTNNB1 (92kDa)
GAPDH (37kDa)
KRJ1
(GEP Neuroendocrine)
B
Pracinostat
Pracinostat
DMSO
UHRF1 (95kDa)
GAPDH (37kDa)
HF2597
(Glioblastoma)
C
Pimasertib
Pimasertib
DMSO
MYC (57kDa)
ACTB (42kDa)
HCC1143
(Breast)
D
Amsacrine
Amsacrine
DMSO
CENPF (330kDa)
ACTB (42kDa)
HF2597
(Glioblastoma)
E
CENPF
merged
GFP
DAPI
Amsacrine
Bortezomib
DMSO
U87
(Glioblastoma)
F
MYC
merged
GFP
DAPI
Pimasertib
Refametinib
DMSO
LoVo
(Colon)
G
qRT-PCR Results
TF / cell line / drug
UHRF1 / H1793 / Bortezomib
CENPF / HF2597 / Bortezomib
UHRF1 / H1793 / Pracinostat
UHRF1 / HF2597 / Bortezomib
Relative expression level compared to DMSO
1.00
0.75
0.50
0.25
0.00
A B C A B C A B C A B C
Three biological replicates for each test