How do small molecules exert their effects in mammalian cells? This seemingly simple question continues to represent one of the fundamental challenges of modern translational science and as such has long been the subject of intense scientific scrutiny. In their recent study, Garnett and colleagues (Gonçalves et al., 2020) demonstrate proof-of-concept for a new way to attack this problem systematically for Oncology drugs, by identifying correlated CRISPR- and drug-killing profiles in the Cancer Dependency Map dataset.

See also: E Gonçalves et al (2020)

Deciphering the mechanism(s)-of-action (MoA) by which small molecules act in human cells is key to identifying why patients do or do not respond to treatment, for developing next-generation molecules with improved efficacy and selectivity and for identifying and preempting mechanisms of resistance. Yet, doing so is highly challenging.

For small molecules emerging from target-based discovery campaigns, it is assumed that the answer revolves around the biochemical perturbation of the target. Despite this assumption, polypharmacological effects may play significant roles. Reciprocally, small molecules that emerge from phenotypic cellular-based screens have historically required laborious biophysical approaches to unmask their MoA. A highly cited example is the elegant discovery of the MoA of the natural product rapamycin: the FKBP12 protein that disrupts mTOR signaling (Brown et al., 1994).

In the case of cancer, several promising new genomic frontiers are now emerging that are beginning to accelerate progress in MoA. First, the use of genetic RNAi or CRISPR modifier screens to identify rescue or sensitization to anti-cancer drug killing has been a powerful approach (Jost & Weissman, 2018; Colic et al., 2019). Second, the use of gene expression and/or high content imaging as a surrogate measurement has enabled the assessment of “connectivity” in signature space (e.g., between a known perturbation and that of a small molecule) (Subramanian et al., 2017). Despite these advances, such approaches are typically limited to specific cancer contexts.

But can the power of CRISPR be leveraged to resolve the MoA of a small molecule for cancer systematically across a wide diversity of cellular contexts? This question is the focus of Gonçalves et al., reported in the current issue of Molecular Systems Biology. The authors hypothesize that a small molecule and a CRISPR genetic knockout that exert the same pattern of killing across cancer cell line models are likely to function through similar mechanisms.

To address this hypothesis, the authors leverage recently emerging data from the Cancer Dependency Map (depmap.org; dmpmap.sanger.ac.uk) in which hundreds of molecularly characterized cancer cell line models have been similarly subjected to genome-wide RNAi or CRISPR (Tsherniak et al., 2017; Behan et al., 2019) and pharmacological (Iorio et al., 2016) profiling. These data have recently been found to be highly reproducible across institutions, suggesting an opportunity for integration to increase power.

The premise in this proof-of-concept manuscript that focuses on established cancer drugs with largely known MoAs is that the correlation in viability between one of ~17,000 genetic knockouts and one of 397 established drugs across 484 diverse cell lines should rediscover the MoA. While examples of success have been reported, a broad-scale study of this new use of the Cancer Dependency Map data has only become possible recently.

Through an extensive series of supervised linear regression analyses, they demonstrate the merits of this approach. They find that in 26% of cases, the killing pattern of the drug is directly phenocopied by the CRISPR killing pattern of the known drug target. They investigate the 264 cases in which this is not the case and find, using protein–protein interactions, 76 additional examples in which the small molecule’s killing pattern correlates with a first, second, or third order interactor. Thus, in aggregate, the authors conclude that either the target or pathway can be rediscovered in roughly 48% of cases.

The authors next investigate some interesting examples. For instance, in the example of isoform-specific PI3K inhibitors, the expected genetic knockdowns indeed correlate. In the case of MCL1 inhibition, they discover an exciting relationship between the MARCH5 E3 ligase and MCL1 inhibition. This finding has now been confirmed by other groups including reports of the MARCH5-dependent degradation of the MCL1/NOXA complex. This discovery should prove very interesting for the therapeutic exploitation of the MCL1 dependency in many human cancers.

Finally, the authors explore concordant biomarkers that explain the sensitivity of both genetic and small molecule perturbations. They show how these biomarkers can
inform MoA, along the lines of previous work focused on small molecules (Rees et al., 2015). For instance, the authors identify tumor necrosis alpha expression as a robust biomarker of sensitivity to both cIAP inhibitors as well as genetic knockout of members of the cIAP pathway.

This manuscript shares some similarities to a recent report (Corsello et al., 2020) that together lay out the exciting potential to resolve the MoA of small molecules in a new way, using the Cancer Dependency Map (Fig 1). One limitation in the current report is the focus on established anti-cancer drugs, which have potent efficacy and (typically) highly refined MoA. Thus, while this is a useful proof-of-concept, it is unclear ultimately how the approach will work where the real MoA challenge is, which is for compounds in development, in which the drug-killing effect is weaker and often less specific.

Additionally, the regression methods used in this paper may not fully account for the notion that most small molecules have diverse polypharmacological effects that together may conspire to explain the
mechanism of killing across various cellular contexts. As the scale of the Dependency Map dataset grows, future computational approaches may be leveraged to produce consensus matches that aggregate individual targets in more complex ways.

With these caveats aside, this report provides new insight into how valuable the Cancer Dependency Map is likely to be as a reference to guide MoA studies throughout the drug discovery process (Fig 1). In the years ahead, it will be important for the field to aspire to learn generalizable lessons around the strengths and weaknesses of various experimental and computational modalities to finally solve this important challenge for oncology once and for all.

References


License: This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
Discovering the anticancer potential of non-oncology drugs by systematic viability profiling


Anticancer uses of non-oncology drugs have occasionally been found, but such discoveries have been serendipitous. We sought to create a public resource containing the growth–inhibitory activity of 4,518 drugs tested across 578 human cancer cell lines. We used PRISM (profiling relative inhibition simultaneously in mixtures), a molecular barcoding method, to screen drugs against cell lines in pools. An unexpectedly large number of non-oncology drugs selectively inhibited subsets of cancer cell lines in a manner predictable from the molecular features of the cell lines. Our findings include compounds that killed by inducing phosphodiesterase 3A-Schlafen 12 complex formation, vanadium-containing compounds whose killing depended on the sulfate transporter SLC26A2, the alcohol dependence drug disulfiram, which killed cells with low expression of metallothioneins, and the anti-inflammatory drug tepoxalin, which killed via the multidrug resistance protein ATP-binding cassette subfamily B member 1 (ABCB1). The PRISM drug repurposing resource (https://depmap.org/repurposing) is a starting point to develop new oncology therapeutics, and more rarely, for potential direct clinical translation.

The prospect of repurposing existing drugs for new clinical indications is alluring: rapid clinical translation can occur for drugs already proven safe in humans. In principle, existing drugs can also establish starting points for drug development when new targets of old drugs are discovered. To date, most oncology repurposing discoveries have been serendipitous; systematic, a-scale screening of the entire pharmacopoeia has not been feasible. The extent to which non-oncology drugs have potential as future cancer therapeutics is unknown.

However, recent efforts have demonstrated the power of large-scale cancer cell line screening—testing either many compounds across a limited number of cell lines (for example, the NCI-60 panel) or a modest number of oncology compounds across many cell lines (for example, the Genomics of Drug Sensitivity in Cancer (GDSC) project at the Sanger Institute and the Cancer Target Discovery and Development (CTD) project at the Broad Institute; Fig. 1a). The ideal study would involve screening many drugs (most of which are non-oncology drugs) across a large panel of genomically characterized cell lines to capture the molecular diversity of human cancer.

In this study, we report the feasibility of using the PRISM molecular barcoding and multiplexed screening method to test 4,518 existing drugs against 578 cancer cell lines. We find that non-oncology drugs have an unexpectedly high rate of anticancer activity. The sensitivity of cancer cell lines to many of these compounds can be predicted from the genomic features of the cell lines, thereby suggesting potentially relevant patient populations.

Results

Drug selection and PRISM profiling. To facilitate the screening of thousands of compounds across hundreds of cell lines, we used the PRISM method. Cancer cell lines are labeled with unique DNA relative barcode abundance serving as a surrogate for cellular viability (Fig. 1b). We screened 578 adherent cell lines spanning 24 tumor types (Extended Data Fig. 1a and Supplementary Table 1). We chose 4,518 drugs from the Drug Repurposing Hub (https://www.broadinstitute.org/repurposing) and confirmed the identity and purity of all compounds to be greater than 75% pure by liquid chromatography–mass spectrometry (LC–MS) (Supplementary Table 2); 3,350 of the compounds (74%) are either approved for clinical use in the USA or Europe, or are in clinical development. The remaining 1,168 (26%) are tool compounds with known activities.

Most compounds, 3,466 (77%), were non-oncology-related, with
the remaining compounds being either chemotherapeutics (2%) or targeted oncology agents (21%).

Screening results. We employed a 2-stage screening strategy whereby drugs were first screened in triplicate at a single dose (2.5 μM); 1,448 drugs screening positives were then rescreened in triplicate in an eight-point dose–response ranging from 10 μM to 610 pM (Fig. 1c and Supplementary Table 2). Interestingly, most active compounds (774 out of 1,448, 53%) were originally developed for non-oncology clinical indications (Fig. 1d). The primary and secondary screening datasets are available on the Cancer Dependency Map portal (https://depmap.org/repurposing) and figshare (https://doi.org/10.6084/m9.figshare.9393293; Extended Data Figs. 1–4). We compared the PRISM results to two gold standard datasets: GDSC (ref. 2) and CTD² (ref. 3). The three datasets revealed similarly high concordance to GDSC and CTD² (Pearson correlations of 0.60 and 0.61, respectively over all shared data points), as the GDSC and CTD² datasets had to each other (Pearson correlation 0.62) yielding 16,650 shared data points. The PRISM dataset had a similar degree of concordance to GDSC and CTD² (Pearson correlations of 0.60 and 0.61, respectively over all shared data points), as the GDSC and CTD² datasets had to each other (Pearson correlation 0.62) (Extended Data Fig. 5a). The three datasets remained similarly concordant when the analysis was restricted to data points showing evidence of anticancer activity (Extended Data Fig. 5b). We conclude that, despite differences in assay format, sources of compounds⁵ and sources of cell lines⁶, the PRISM Repurposing dataset is similarly robust compared to existing pharmacogenomic datasets.

At the level of individual compound dose–response, we note that the PRISM Repurposing dataset tends to be somewhat noisier, with a higher standard error estimated from vehicle control measurements (Extended Data Fig. 5c and Extended Data Fig. 6a–c). This variation may be explained by a combination of longer assay duration, smaller number of cells assayed and/or variation attributable to growing cells in pools. However, such noise was not substantial enough to preclude the discovery of anticancer activities or their associated predictive biomarkers (see further on).

Landscape of non-oncology drug effects on cancer viability. We performed unsupervised clustering of compound viability profiles independent of their functional annotations using the uniform manifold approximation and projection (UMAP) method⁷ (Fig. 2a; interactive plot available at https://depmap.org/repurposing). Compounds with similar mechanisms of action tended to cluster together, indicating that expected activities were recovered by the PRISM assay. Interestingly, while we expected to recover known mechanisms of action for cancer drugs, we also found clusters of functionally related noncancer drugs, such as vitamin D receptor agonists and 3-hydroxy-3-methylglutaryl-coenzyme A reductase inhibitors. Of note, some functionally related classes (for example, glucocorticoid receptor agonists) showed two or more distinct clusters, suggesting that biologically relevant substructures may exist within the dataset.

In general, chemotherapeutics killed the highest number of cell lines and non-oncology drugs the lowest, with targeted oncology drugs being intermediate (Fig. 2b). However, this pattern was highly dose dependent. At high doses, targeted agents lost their selectivity. Perhaps most interestingly, a subset of non-oncology drugs showed particularly potent activity: 91 drugs killed at least 1% of cell lines at a concentration of 625 nM or lower.
To further investigate the therapeutic potential of non-oncology drugs, we computed the bimodality coefficient of each compound’s dose-wise viability profile (Extended Data Fig. 6d and Supplementary Table 3) and then calculated the maximum for each compound. While non-oncology drugs showed less bimodality than cancer drugs on average (Fig. 2c), the most selective compounds in...
the dataset include non-oncology drugs (Fig. 2d). This provides further evidence that large-scale testing of noncancer drugs can reveal selective anticancer activity.

**Predictive models of killing activity.** We next addressed the extent to which cell-killing activity was predictable based on the genomic features of cell lines. For each drug, we used the random forest-based ATLANTIS algorithm, which employs the baseline molecular features (cell lineage; gene copy number; function-damaging, hotspot or missense mutations; DNA methylation levels; messenger RNA, microRNA, protein or metabolite abundance) and genetic dependencies (genome-scale CRISPR–Cas9 knockout or RNA interference) as inputs to the model.10–14 Most of the highly predictable killing profiles came from targeted oncology drugs, compared to chemotherapeutics or non-oncology drugs (Fig. 2e and Supplementary Tables 3 and 4). This is not surprising, since targeted oncology drugs have been optimized for their selective killing of subtypes of cancer. More striking was the observation that a substantial number of non-oncology drugs had highly predictable patterns of killing (38 with a Pearson score >0.4 in the secondary screen; Fig. 2f).

Interestingly, mRNA expression was by far the most predictive feature type compared to other categories of genomic information (Fig. 3a). This observation is consistent with the findings of other pharmacogenomic screens15–18 as well as short hairpin RNA-based functional genomic screens19. By contrast, Iorio et al.17 reported that mutation, not gene expression, was most predictive of drug response within a cancer type. Whether that is explained by their focus on known cancer drugs (for example, kinase inhibitors), as opposed to all classes of drugs analyzed in the present dataset, is to be determined. Importantly, only rarely (in 0.8% of cases) did the pattern of killing by active non-oncology drugs correlate with knockout or knockdown of the drug’s intended target (versus 15.0% of active oncology drugs; Fig. 3b and Supplementary Table 6). This suggests that the unexpected anticancer activity of non-oncology drugs is most probably explained by a previously unrecognized mechanism of action.

In the primary screen, 195 non-oncology drugs had predictable killing with a Pearson score >0.2, and 23 were predictable with a Pearson score >0.4. This same phenomenon of highly predictable non-oncology drug killing was observed when predictability was plotted against the bimodality score (Fig. 4). The most predictable, bimodal non-oncology drugs (right plot; upper right-hand quadrant; Supplementary Tables 3 and 4) represent the drugs of greatest interest for future mechanistic follow-up. We describe four of these compounds in the sections that follow.  

**Inducers of PDE3A–SLFN12 protein–protein interaction.** Among the genomic features most highly correlated with non-oncology drug activity was the expression of the gene PDE3A, whose expression correlated with killing by 11 structurally diverse compounds. These included the known phosphodiesterase 3A (PDE3A) inhibitors anagrelide and zardaverine, progesterone receptor agonists (including the nonsteroidal drug tanaprost), the kinase inhibitor AG-1296 and the potassium channel activator DCEBIO (Fig. 5a). This pattern of killing was of interest because of the recent report of cancer cytotoxicity occurring as a result of protein–protein interaction between PDE3A and the largely uncharacterized protein Schlafen 12 (SLFN12; ref. 19)). We found that the structurally diverse compounds identified in the PRISM screen bound PDE3A in a thermal shift assay (Fig. 5b and Supplementary Table 7) and inhibited PDE3A enzymatic activity (Supplementary Table 7). Their cytotoxicity was completely rescuable by either PDE3A knockout or by competition with trequinsin, a potent PDE3A small-molecule inhibitor that does not induce PDE3A–SLFN12 interaction18 (Fig. 5c).

Importantly, PDE3A pull-down resulted in coimmunoprecipitation of V5-tagged SLFN12 following compound treatment, indicating that these compounds indeed induced PDE3A–SLFN12 protein–protein interaction (Fig. 5d). We consistently observed that, while complex formation predicts compound sensitivity, the strength of the PDE3A–SLFN12 interaction does not directly correlate with HeLa cell line half maximal inhibitory concentration (IC50; Fig. 5d and Supplementary Table 7). Taken together, the PRISM results show that an unexpectedly large number of structurally
and MT2A are known metal-chelating proteins. Consistent with MT2A on 16q are predictive of disulfiram activity is mechanistically plausible: disulfiram's activity is copper-dependent, and MT1E and MT2A are known metal-chelating proteins. Consistent with this observation, MT1E and MT2A expression was also correlated with sensitivity to thiram and elesclomol, other copper-binding compounds. In addition, disulfiram has been reported to induce metallothionein gene expression in prostate cancer cells.

To test the hypothesis that metallothionein expression regulates disulfiram's anticancer activity, we used the disulfiram-resistant glioma cell line SF295, which has an amplification of chromosome 16q and high metallothionein expression. To inhibit the expression of multiple metallothionein genes simultaneously, we knocked out the transcription factor MTF1, which is a known upstream regulator of metallothionein gene expression (Extended Data Fig. 7a; ref. 22). Following MTF1 knockout, metallothionein genes were downregulated as assessed by global mRNA sequencing (Extended Data Fig. 7b). As predicted, MTF1 knockout resulted in increased sensitivity of SF295 cells to disulfiram; this increased sensitivity could be completely reversed by the copper chelator tetrathiomolybdate (TTM; Fig. 6b). MTF1 knockout did not alter sensitivity to control bortezomib (Extended Data Fig. 7c). These results together suggest that 16q deletion is a potential predictive biomarker of disulfiram and other copper-dependent cytotoxic agents. This finding is clinically relevant because arm-level 16q loss is seen in many tumor types, most notably breast and ovarian cancers, where its prevalence is estimated at 55–65% and 55–76%, respectively.

Vanadium-containing compounds. The PRISM screen revealed a strong correlation between killing by the vanadium-containing drug bis(maltolato)oxovanadium(IV) (BMOV) and expression of the sulfate transporter SLC26A2 (Pearson correlation −0.583, ATLANTIS Pearson score 0.620; Fig. 6c). Bioavailable vanadium-containing compounds, including BMOV and bis(ethylmaltolato)oxovanadium(IV) (BEOV), have been of interest for their ability to lower fasting blood glucose in animal models and in patients with diabetes. SLC26A2 function has not been extensively studied, but loss-of-function mutations have been associated with connective tissue disorders. In cancer, SLC26A2 is broadly expressed at modest levels, with high expression in melanoma and uterine cancers.

The mechanistic relationship between BMOV and SLC26A2 expression is not obvious. Arguing against BMOV being an inhibitor of SLC26A2 function, analysis of publicly available genome-wide CRISPR–Cas9 loss-of-function screens indicates that SLC26A2
is not a cancer dependency (see https://depmap.org). Consistent with this, knockout of SLC26A2 in the BMOV-sensitive cell line OVISE was tolerated (Extended Data Fig. 7d,e). However, SLC26A2 knockout rendered OVISE cells resistant to both BMOV and BEOV, indicating that SLC26A2 is not simply a biomarker of killing, but is required for compound activity (Fig. 6d). The cytotoxicity of other, structurally distinct vanadium-containing compounds was similarly rescued by SLC26A2 knockout, suggesting that the vanadium oxide ion is responsible for the SLC26A2-dependent cytotoxicity of BMOV (Fig. 6e). It remains to be determined whether these compounds are substrates for the SLC26A2 transporter, interfere with sulfate ion homeostasis or confer a previously unknown function to SLC26A2.

**Tepoxalin and multidrug resistance.** Expression of metabolic enzymes or drug efflux pumps were among the most common predictive biomarkers of drug response in the PRISM screen. As expected, high mRNA expression of the gene encoding the ABCB1 transporter (also known as multidrug resistance protein 1 (MDR1) or p-glycoprotein 1) was the top predictor of resistance to numerous
oncology drugs, including taxanes (docetaxel and paclitaxel), vinca alkaloids (vincristine and vinorelbine) and proteasome inhibitors (carfilzomib) (Fig. 7a).

An unexpected finding, however, was that a single drug, tepoxalin, had the opposite relationship to ABCB1: high ABCB1 expression predicted sensitivity to tepoxalin. Tepoxalin is a dual cyclooxygenase

Fig. 6 | Anticancer activity of disulfiram and vanadium. a, Disulfiram sensitivity of PRN,N’-Bis(salicylidene)phenylenediamine vanadium(IV) oxide(ISM cell lines grouped by 16q copy number status (n = 447 cell lines). PRISM secondary data are shown as AUCs. Arm-level copy number data were obtained using published The Cancer Genome Atlas methods and manually reviewed to ensure consistency with copy number at the 16q13 locus. Two-sided P values were calculated using the Wilcoxon signed-rank test between each pair of groups. The upper box limits, center lines and lower box limits correspond to the 75th, 50th and 25th percentiles, respectively. The whiskers extend from the box limits to the most extreme value up to 1.5 interquartile range from the median. All cell lines are depicted as points, regardless of outlier status. b, MTF1 loss sensitizes to disulfiram. Drug sensitivity of SF295 cells with and without MTF1 knockout, with sgRNA against GFP included as a nontargeting control. Cotreatment with 1 μM TTM is included. Mean viability across three independently treated wells is shown, with the s.d. indicated by the error bars. c, Scatter plot of BMOV sensitivity versus SLC26A2 gene expression (n = 538 cell lines). Correlation between PRISM viability data at 2.5 μM and SLC26A2 gene expression from CCLE RNA-seq (log, RPKM) is shown. Pearson’s r is shown. d, SLC26A2 knockout confers resistance. Dose-response curves for BMOV and BEOV in OVISE cells with and without CRISPR–Cas9-mediated knockout of SLC26A2. sgRNA against GFP was included as a nontargeting control. Mean viability across three independently treated wells is shown with the s.d. indicated by the error bars. e, Dose-response curves for related vanadium-containing compounds, N,N’-Bis(salicylidene)-o-phenylenediamine vanadium(IV) oxide and vanadyl sulfate, in OVISE cells with and without SLC26A2 knockout. Mean viability across three independently treated wells is shown, with the s.d. indicated by the error bars.
and 5-lipoxygenase inhibitor that is Food and Drug Administration-approved for the treatment of osteoarthritis in dogs\textsuperscript{14,15}. Because more than 100 other cyclooxygenase and/or 5-lipoxygenase inhibitors were also tested in our PRISM screen, we asked whether they shared the ABCB1-associated killing effect. Strikingly, none of them did, suggesting that tepoxalin’s killing was most probably explained by an off-target mechanism. Consistent with this hypothesis, tepoxalin’s killing activity was not correlated with the genetic knockout profiles of its known targets PTGS1, PTGS2 and ALOX5.

To gain insight into the mechanism by which tepoxalin selectively inhibits cancer cells, we performed a genome-wide CRISPR–Cas9 modifier screen to identify the genes required for tepoxalin-mediated activity. LS1034 colorectal cancer cells have high levels of ABCB1 expression and are inhibited by tepoxalin with an IC\textsubscript{50} of 3.8 µM. Cas9-expressing LS1034 cells were infected with a pooled library containing 76,441 single guide RNAs (sgRNAs) targeting 19,114 genes and treated with 16 µM tepoxalin (or vehicle control) for 28 d. Remarkably, the gene knockout most enriched in tepoxalin-resistant cells was ABCB1 itself (Fig. 7b). Other resistance hits included multiple components of the SWI/SNF complex (SMARCA4, SMARCB1 and ARID1A), which has been previously implicated in the regulation of ABCB1 gene expression\textsuperscript{16}.

**Fig. 7** | Tepoxalin is active against ABCB1-high cancer cell lines via an ABCB1-mediated mechanism. **a** Drug sensitivity profiles of tepoxalin and seven additional compounds where ABCB1 RNA expression was the top predictive genomic feature of the PRISM activity profile (n = 489 cell lines). Tepoxalin was the only compound tested where high ABCB1 expression was associated with sensitivity rather than resistance to tepoxalin. Cell line viability is depicted with ABCB1 gene expression from CCLE RNA-seq log\textsubscript{2} RPKM data. Cell lines are ranked by mean viability with the ceiling at 100%. **b** ABCB1 knockdown is the top hit that rescues tepoxalin activity in a genome-wide CRISPR–Cas9 gene knockout screen. LS1034-Cas9 cells (pXPR_311) were infected with the Brunello sgRNA library, selected with puromycin and treated with 16 µM tepoxalin versus vehicle control with two replicates in independent flasks. Cells were passaged every 3–4 d over a 30-d period. Two-sided \( P \) values were computed using MAGeCK-MLE and are shown versus gene-level log fold changes. **c** ABCB1 overexpression sensitizes to tepoxalin activity in a genome-wide CRISPR–Cas9 gene activation screen. LS1034-dCas9 cells (pXPR_109) were infected with the Calabrese sgRNA library, selected with puromycin and treated with 16 µM tepoxalin versus vehicle control with two replicates in independent flasks. Cells were passaged every 3–4 d over a 14-d period. Two-sided \( P \) values were computed using MAGeCK-MLE and are shown versus gene-level log fold changes. **d** Tepoxalin cellular competition assay after ABCB1 knockout. LS1034-Cas9 cells (pXPR_311) were stably infected with firefly luciferase; parental LS1034 cells (without Cas9) were stably infected with Renilla luciferase. Cells were mixed in a 1:1 ratio and infected with the indicated sgRNA construct against ABCB1 or an intergenic region on chromosome 2 (negative control). After puromycin selection, cell mixtures were treated with 16 µM tepoxalin. The firefly:Renilla luminescence ratio is plotted as the log fold change over time. The mean of three technical replicates is shown and the results are representative of three independent experiments.
To complement the CRISPR–Cas9 loss-of-function screen, we performed a genome-wide CRISPR activation screen to identify genes whose overexpression was selected against in the setting of tepoxalin treatment. The gene showing the most negative selection was also ABCB1, indicating that its overexpression sensitizes cells to tepoxalin (Fig. 7c). Consistent with these findings, cellular competition and dose–response assays revealed strong selection for tepoxalin-treated ABCB1-null cells compared to wild-type cells (Fig. 7d, Extended Data Fig. 7f–h and Supplementary Table 8). Overexpression of ABCB1 in a low-expressing, tepoxalin-insensitive cell line resulted in increased drug sensitivity (Extended Data Fig. 8a,b). Furthermore, potent ABCB1 small-molecule inhibitors did not phenocopy tepoxalin in the PRISM assay and in fact antagonized tepoxalin-induced killing, suggesting that ABCB1 inhibition alone does not explain tepoxalin’s anticancer activity (Extended Data Fig. 8c,d).

Given that ABCB1 encodes a drug transporter, we next asked whether ABCB1 expression affected intracellular concentrations of tepoxalin. However, LC–MS experiments indicated that intracellular concentrations of tepoxalin were unaffected by levels of ABCB1 expression or by ABCB1 small-molecule inhibition (Extended Data Fig. 8e). Tepoxalin is also known to be metabolized to a compound known as RWJ20142 by conversion of its hydroxamic acid to a carboxylic acid42. RWJ20142 is generated in the presence of serum but not saline or acetonitrile (Extended Data Fig. 8f). Unlike tepoxalin, RWJ20142 was not cell permeable and showed no inhibition of ABCB1-expressing cancer cells. Although tepoxalin inhibited ABCB1 at high concentrations (Extended Data Fig. 8g,h), our results indicate that tepoxalin does not kill cancer cells simply by inhibiting ABCB1 activity. Taken together, these results suggest that tepoxalin, but not its metabolite RWJ20142, inhibits ABCB1 high-expressing cancer cells via an ABCB1-mediated mechanism that remains to be fully elucidated.

Discussion

We developed the PRISM Repurposing dataset as a large-scale resource containing the anticancer activity of non-oncology drugs. The PRISM screen recovered 49 non-oncology compounds with selective and predictive biomarker-associated anticancer activity (Pearson score > 0.2, bimodality coefficient > 0.4) and 103 with a less stringent bimodality coefficient cutoff (>0.35). Of note, six non-oncology compounds (Fig. 4) showed selective (bimodal) killing patterns, but their activity was unpredictable based on the baseline genomic features of the cell lines. It is possible that an expansion of the number of cell lines in the PRISM panel will help identify such biomarkers. Alternatively, killing might be explained by molecular features not yet measured in the cell lines.

It is conceivable that some non-oncology drugs could be brought directly to clinical trials for testing in cancer patients. However, before doing so, it is important to establish that the killing activity of such drugs is observed at concentrations that are achievable and tolerable in humans. Similarly, it is important to confirm that the predictive biomarkers identified in cell lines represent distinct populations of human tumors in vivo. It is probable that most of the observations described in this study will not be suitable for immediate testing in humans; either the biomarker hypothesis will require further refinement or the compounds themselves will require further optimization.

In contrast to immediate repositioning of existing drugs for new indications, the PRISM results reported in this study also represent starting points for new drug development. When the anticancer activity of a drug occurs via an off-target mechanism, it is probable that further optimization for this new target will result in more potent and selective drug candidates. We note that the use of cell-based screens such as PRISM allows for the discovery of previously unknown mechanisms of action that would be difficult to discover using conventional biochemical screening assays.

We initiated follow-up studies on four of the initial findings from the PRISM screen. In the case of disulfiram, we discovered a previously unrecognized biomarker (16q deletion) that predicts sensitivity. Future work will require extending such studies to the in vivo setting and determining whether sufficiently high disulfiram concentrations can be achieved to obtain anticancer effects. The accessibility of copper in different organs and cell types probably also modulates disulfiram’s anticancer activity. Compounds with remarkable chemical diversity kill PDE3A high-expressing cancer cell lines. They all induced PDE3A–SLFN12 complex formation, as has been described for the compound DNMDP48. A structural understanding of the interaction of these diverse compounds with PDE3A will probably inform future optimization of PDE3A–SLFN12-directed cancer therapeutics. Our finding that vanadium-containing compounds selectively kill cancer cells expressing high levels of the sulfate transporter SLC26A2 was also surprising, given that a mechanistic link between the two had not been previously suspected. A recent study showed that SLC26A2 expression is a mechanism of resistance to TNF-related apoptosis-inducing ligand-induced cell death49; it remains to be determined whether that mechanism is relevant to vanadium-induced killing and whether these compounds dysregulate sulfate homeostasis. Perhaps most interesting was our observation that the drug tepoxalin has the unique ability to inhibit cells that express high levels of the multidrug resistance gene ABCB1. While tepoxalin was originally developed as a cyclooxygenase/5-lipoxygenase inhibitor, our structure–activity relationship studies clearly showed that tepoxalin’s anticancer activity is probably cyclooxygenase- and 5-lipoxygenase-independent. While we showed that ABCB1 is both necessary and sufficient to confer tepoxalin cytotoxicity, the precise mechanism by which such cell death occurs is to be established. Further optimization of tepoxalin against this new target, and engineering out the drug’s cyclooxygenase/5-lipoxygenase inhibitory activity, would probably result in improved tolerability as an anticancer agent.

Historically, a challenge with cell-based phenotypic screens is the difficulty in gaining molecular insight into the mechanism of action of hit compounds. However, our results in the present study demonstrate the power of genome-scale CRISPR–Cas9 loss-of-function and gain-of-function screens to provide mechanistic clues to small-molecule action. For example, both CRISPR knockout and CRISPR activation screens pointed to ABCB1 as the most relevant target of tepoxalin. The availability of such functional genomic screening methods will probably reinvigorate cell-based screening more broadly. We also note that, whereas small molecules are typically thought of as inhibitors of their protein targets, our PRISM results indicate that this is often not the case. For example, we discovered compounds that stabilize protein–protein interaction (for example, PDE3A–SLFN12) and that engage ABCB1 but do not kill cells by ABCB1 inhibition. A plethora of noninhibitory small-molecule activities probably remain to be discovered from this PRISM dataset. However, we note that when anticancer drug targets are not themselves cancer dependencies, it is possible that strong selective pressure will result in downregulation of the target protein, resulting in drug resistance. Such drug resistance could presumably be overcome through combination drug treatment, as is the norm for most types of cancer.

The PRISM barcoding and pooling approach described in this study substantially increases screening efficiency, but it is conceivable that the pooling of cell lines results in paracrine-mediated mechanisms that modulate drug sensitivity. In practice, however, we have yet to observe such cell–cell interactions or any consistent discordance between PRISM and one-by-one viability profiling. Nevertheless, we and others have reported the existence of microenvironment-mediated drug resistance mechanisms50; the potential for such interactions should be considered when interpreting the PRISM results.
The PRISM Repurposing dataset described in this study represents nearly half of all drugs ever tested in humans. Given the large number of unexpected findings that emerged from this initial screen, we believe that expansion of the PRISM resource in both the dimension of drugs and cancer models is warranted. Such data will provide an important pharmacological component of the Cancer Dependency Map (https://depmap.org), which in turn will form a preclinical foundation for cancer precision medicine.

Methods

Cell lines. Parental cell lines were obtained from the Broad Institute-Novartis Institutes for BioMedical Research Cancer Cell Line Encyclopedia (CCLE) project before PRISM barcoding (see https://portals.broadinstitute.org/ccle for the original sources). For the follow-up studies, LS1034, HeLa and HEK293T cells were purchased from the ATCC. The REC1, SF295, OVISe, COLO 320, HT-29 and SNU-449 cell lines were provided by the Cancer Dependency Map (Broad Institute). Short tandem repeat (STR) fingerprinting was performed by Genetica using the PowerPlex 16 HS system (Promega Corporation). STR profiles were compared with STR profiles reported by vendors and in literature. Misidentified cell lines or other STR conflicts are listed in Supplementary Table 9. These cell lines are flagged in the data files and are not shown by default in the website interface. Cell lines were confirmed to be negative for Mycoplasma using the MycoScope PCR Mycoplasma Detection Kit (Genelantis). The LS1034, REC1, OVISe, SF295, COLO 320, SNU-449 and Kuramochi cell lines were cultured in Roswell Park Memorial Institute (RPMI) medium (Thermo Fisher Scientific). HeLa cells were cultured using DMEM (Thermo Fisher Scientific). The LS1034, REC1, SF295, COLO 320, SNU-449 and OVISe cell lines derived with Cas9 were provided by the Cancer Dependency Map (Broad Institute).

PRISM screening. We made several improvements to the PRISM barcoding method described previously. The assay employs a 24-nucleotide barcode stably introduced into cancer cell lines via lenti viral transduction. The barcode is located at the end of the blasticidin resistance gene and is expressed as an mRNA under a tetracycline-inducible promoter. The REC1, SF295, OVISe, COLO 320, HT-29 and SNU-449 cell lines were purchased from the ATCC. The REC1, SF295, OVISe, COLO 320, SNU-449 and OVISe cell lines derived with Cas9 were provided by the Cancer Dependency Map (Broad Institute). Short tandem repeat (STR) fingerprinting was performed by Genetica using the PowerPlex 16 HS system (Promega Corporation). STR profiles were compared with STR profiles reported by vendors and in literature. Misidentified cell lines or other STR conflicts are listed in Supplementary Table 9. These cell lines are flagged in the data files and are not shown by default in the website interface. Cell lines were confirmed to be negative for Mycoplasma using the MycoScope PCR Mycoplasma Detection Kit (Genelantis). The LS1034, REC1, OVISe, SF295, COLO 320, SNU-449 and Kuramochi cell lines were cultured in Roswell Park Memorial Institute (RPMI) medium (Thermo Fisher Scientific). HeLa cells were cultured using DMEM (Thermo Fisher Scientific). All media were supplemented with 10% heat-inactivated FCS (Sigma-Aldrich) and 1% penicillin-streptomycin (Thermo Fisher Scientific). HeLa cells were cultured using DMEM (Thermo Fisher Scientific). All media were supplemented with 10% heat-inactivated FCS (Sigma-Aldrich) and 1% penicillin-streptomycin (Thermo Fisher Scientific) except for the HEK293T cell line, which was maintained without antibiotics.

Data processing. Luminex median fluorescence intensity (MFI) values were calculated as the median fluorescence values of all beads corresponding to a single PRISM barcode in a single technical replicate. MFI values were log-transformed (logMFI) and subjected to two quality control steps. First, an outlier pool filter was applied to remove probable screening artifacts (Extended Data Fig. 2). In each assay plate, logMFI values were median centered per cell line and each well was summarized by the median of these centered MFI values. Wells more than five median absolute deviations from the median across all wells from the same compound plate and plate location were removed. Second, a ‘control separation filter’ was applied. For each plate, cell lines with strictly standardized mean difference values <2 were excluded from the rest of the analysis (Extended Data Fig. 3). Strictly standardized mean difference values were calculated as:

$$\frac{\mu_r - \mu_s}{\sqrt{\sigma_r^2 + \sigma_s^2}}$$

where $\mu_r$ and $\sigma_r$ stand for the medians and the median absolute deviations of the logMFI values computed over the negative/positive control wells for each cell line on each plate. The number of PRISM profile cell line replicates passing quality control is shown in Extended Data Fig. 4.

A total of 1,448 compounds were selected for secondary eight-point dose–response testing based on reproducibility, predictability, selectivity and compound availability (Supplementary Table 2). For the secondary screen only, ten unique barcodes were spiked into each well of each plate after cell lysis. Normalized MFI values were computed by taking the ratio of each logMFI value to the median logMFI of the inert barcodes in each well. For data produced before the spike-in protocol was introduced, normalized MFI values were set equal to MFI values.

Fold change values were calculated as the ratio of normalized MFI to the median of the normalized MFI from dimethyl sulfoxide (DMSO)-treated negative controls for each cell line on each plate. The batch effects produced from variable assay conditions were then removed using ComBat (located in the sva package; we used version 3.30.1 of sva)45. ComBat was run over each treatment condition separately by considering the log-transformed fold change values as the probes and the pool-replicate combinations as the batches. The corrected log fold change values were then median collapsed for each cell line, screen, source plate and well combination. We labeled cell lines as sensitive to a treatment if the median collapsed fold change was <0.3.

Dose–response analysis. Dose–response relationships were obtained by fitting four-parameter logistic curves to viability values for each cell line using the R package drc (version 3.0-1). Following the practice of Smirnov et al.17, the upper asymptote of the logistic curves was fixed at 1 and the viability values were fitted as a function of drug concentration according to:

$$V(c) = \frac{E_{\text{max}}}{1 + \left(\frac{c}{E_{\text{EC50}}}ight)^{k}}$$

where all concentrations are in the natural logarithm scale. $E_{\text{max}}$, values were defined as the concentration at which $V(c) = 0.5$. Additionally, the dose–response area under the curve (AUC) was calculated using the normalized integral:

$$\text{AUC} = \frac{\int_{c_{\text{min}}}^{c_{\text{max}}} V(c)dc}{\frac{c_{\text{max}} - c_{\text{min}}}{E_{\text{max}} - E_{\text{min}}}}$$

This formulation puts AUC values on a scale between 0 and 1 for curves with lower asymptotes $<1$, where lower AUC values indicate increased sensitivity to treatment.

Biomarker discovery. To generate predictive biomarkers, we adopted the ATLANTIS predictive models’ and trained multiple models for each PRISM profile. ATLANTIS is a tailored nonlinear regression model for gene dependency prediction based on the baseline characteristics of cancer cell lines. More specifically, ATLANTIS is an efficient implementation of a conditional inference forest11 with additional weighting and iterative feature selection steps. Implementation details have been published previously46 and the code is available on a public repository (https://github.com/cancerdatasci/atlantis).

For each dose-wise log fold change profile, 14 ATLANTIS models are trained (1 model per feature set). Feature sets include baseline cell line omics, genetic dependencies and experimental confounders. All feature sets are listed in Supplementary Table 10.

Next, the predictive performance of each model was assessed based on the Pearson correlation between out-of-bag model predictions and the response variable. Models with Pearson correlations greater than 0.2 were strong models. The relative importance of each feature (degree of decrease in accuracy) was computed by ATLANTIS for each model. The most important feature of each strong predictive model is presented as a potential predictive biomarker or strongly associated phenotype. The comprehensive list of biomarkers is available at https://depmap.org/repurposing.

Compound killing selectivity. To assess for selective killing activity, the bimodality coefficient for each median-collapsed log fold change PRISM profile was computed for each compound as follows:

$$g^2 = \frac{n}{k + 3(n - 1)^2/(n - 2)(n - 3)}$$

where $n$ is the number of samples (cell lines), $g$ is the sample skewness and $k$ is the sample excess kurtosis. Note that a larger bimodality coefficient implies a highly skewed (large magnitude of $g$) but light-tailed (small kurtosis) distribution.

Computation of AUC values for cross-dataset comparison. Secondary PRISM Repurposing data were compared to CTD+ (v.2.0, accessed 15 December 2015) and GDSC available through the PharmacoGX package (version 1.12.0)47. For the PRISM Repurposing and GDSC datasets, dose–response curves were fitted as described earlier. CTD+ provides dose–response curve parameters and curves were not refitted. The scope of the comparison was limited to compounds screened in all three datasets with a minimum overlapping fourfold dose range across the datasets. Dose–response curves were computed for each compound-cell line dataset combination using all available doses. AUC values were calculated over the shared dose range (curves were not refitted). The complete table of the
**Assessment of noise in the PRISM Repurposing dataset.** The s.e.m. of the inferred log fold change viability values was calculated to estimate the amount of noise in the PRISM screen. We assumed that (spike-in) normalized logMFI values have a cell line-specific additive noise with a constant variance σ² across treatments. An s.e.m. propagation analysis (see section 9.3 of Chatfield⁵), gives the following: \[ \sigma_{\text{log fold change}}^2 = \sigma_{\text{treatment}}^2 + \sigma_{\text{control}}^2 = \frac{\sigma^2}{1 + 1/N_{\text{control}}} + \frac{\sigma^2}{N_{\text{rep}}} \]

where \(N_{\text{rep}}\) is the number of replicates (3 for PRISM Repurposing), \(N_{\text{control}}\) is the number of negative control wells in a given plate (52 in the standard PRISM assay format) and \(\sigma_{\text{log fold change}}\) is the s.e.m. estimate; σ was estimated separately for each cell line and plate using normalized, negative control log MFI values. σ estimates for the rest of the parameters were set to the defaults provided by the R package umap7. This procedure was applied separately to each PRISM Repurposing screen. In MTS006, DMSO-only plates were used.

**Projection of viability profiles to two dimensions.** The log viability profiles from the primary screen were embedded into a two-dimensional manifold using the UMAP algorithm and were visualized in Fig. 2a. Cell lines that were missing more than 10% of their viability values (failed in the quality-control steps) were removed and the remaining missing values were imputed using the R package FastImputation (version 2.0)⁶. UMAP was applied to the resulting data, using the cosine distance metric and the following configuration parameters: 7 nearest neighbors, 0.5 minimum distance; 2 components; and 200 training epochs. The rest of the parameters were set to the defaults provided by the R package umap (version 0.2.3.1). For the visualization, we filtered out compounds with an average Pearson replicate correlation below 0.25. Mechanisms of action that include fewer than 20 compounds are shown in transparent gray.

**Comparison to genetic loss-of-function screens.** More than 20 compounds are shown in translucent gray. Pearson replicate correlation below 0.25. Mechanisms of action that include fewer than 20 compounds are shown in transparent gray.

**Synthesis of BEOV.** Vanadyl sulfate trihydrate (25 g, 115 mmol) dissolved in 25 ml of water was added to ethyl maltool (43.4 g, 310 mmol) dissolved in 125 ml of hot water under argon; the resulting solution was heated to 100 °C for 30 min. The pH was adjusted very slowly to 8.5 by adding NaOH (12.7 g, 319 mmol) in 10 ml of water. The resulting mixture was refluxed for 2 h and then allowed to cool to room temperature. The dark–blue gray solid was collected by vacuum filtration, washed with cold water and dried in a vacuum to produce BEOV with a yield of 88%. The compound was 93% pure by LC–MS.

**Cloning.** The pXPX_003 and pXPX_023 vectors were acquired from the Genetic Perturbation Platform (GPP, Broad Institute). The oligonucleotides for sgRNA design were generated using Broad GPP sgRNA guide resource generator (https://portals.broadinstitute.org/gpp/public-analysis-tools/sgrna-design) and the oligonucleotide sequences were synthesized by Integrated DNA Technologies. To clone the sgRNAs into either the pXPX_003 'guide-only' or pXPX_023 all-in-one CRISPR lentiviral expression systems, we followed the protocol available on the GPP website (https://portals.broadinstitute.org/gpp/public/resources/protocols). The CRISPR sgRNA sequences are shown in Supplementary Table 12.

**Antibodies and western blot.** The following antibodies were used: polyclonal rabbit anti-PDE3A (1:1,000; catalog no. A302-740A; Bethyl Laboratories); monoclonal mouse anti-V5 (1:5,000; catalog no. R960-25; Thermo Fisher Scientific); monoclonal rabbit anti-MDR1/ABCB1 (clone D3H11Q) (1:1,000; catalog no. 12683; Cell Signaling Technology); monoclonal mouse anti-β-actin (clone 81H1D10) (1:1,000; catalog no. 3700; Cell Signaling Technology); and monoclonal mouse anti-SLC26A2 (clone 26E1) (1:1,000; dilution catalog no. HY-10001386-M04; Novus Biologicals). Cells were lysed with radioimmunoprecipitation assay buffer (CHAPS buffer substituted for Extended Data Fig. 8a) supplemented by protease and phosphatase inhibitors (Sigma–Aldrich). For the ABCB1 blots, proteins were transferred onto a nitrocellulose membrane using a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad Laboratories) in Tris-glycine buffer (Bio-Rad Laboratories) with 10% methanol for 4 h at 60 V and 4°C. Membranes were blocked in Odyssey Blocking Buffer (LI-COR) for 1 h and probed overnight with primary antibodies. The following day, membranes were washed and probed with IRDye secondary antibodies (catalog nos. 926-68020 and 926-32111, each at a 1:5,000 dilution; LI-COR). The FDEA and V5 blots, different secondary antibodies were used (926-32210 and 926-68020, each at a 1:5,000 dilution; LI-COR). Blot images were collected with the Odyssey CLx imager (LI-COR).

**Genomic DNA (gDNA) PCR and next-generation sequencing to quantify CRISPR editing frequency.** To confirm efficient CRISPR cutting at the target loci, PCR primers were designed flanking the sgRNA cut sites by 75–100 base pairs (bp) on either side. The human gDNA PCR primer sequences corresponding to the regions targeted by MDR1, SLC26A2 and MTF1 sgRNAs are shown in Supplementary Table 13. gDNA was isolated from knockout cell lines using the Gentra Puregene Kit (QIAGEN) and amplified with the primers to yield an amplicon roughly 150–250 bp in length using Hercule II Fusion Polymerase (Agilent Technologies). The PCR product involved a 2-min denaturation at 95°C, followed by 24 cycles of 95°C for 2 min, 55°C for 2 min and 72°C for 1 min. PCR samples were purified and submitted for the next-generation sequencing assay at the Massachusetts General Hospital DNA Core. Knockout efficiency was assessed by the percentage of reads containing a frameshift caused by an indel compared to the total read counts. Results for each guide were averaged across primer sets with successful amplification.

**Cellular viability assays.** Cells were seeded at a density of 2,000 cells per well (1,000 cells per well for the HeLa cell line) in a 96-well clear bottom black microplate (Corning). The following day, compounds were treated using the D300e Digital Dispenser (Tecan). Following incubation, viability was assessed by CellTiter-Glo (Promega Corporation). Luminescence was measured using an EnVision plate reader (PerkinElmer; EnVision Manager 1.13.3009.1401).
Independent replicates were washed and normalized to vehicle control. Dose curves were generated using Prism 8 (GraphPad Software).

PDE3A commounocprecipitation assay. The PDE3A immunoprecipitation and western blotting of coprecipitated SLFN12-V5 protein experiments were performed as described previously25. Briefly, HeLa cells were plated onto 15 cm plates at 3 × 10^6 cells per plate and transfected the next day with 15 µg of pLX307-SLFN12 plasmid (clone TRC0000476272) using FuGENE 6 (Promega Corporation) at a 4:1 ratio. Roughly 72h posttransfection, cells were treated with 10µM of the repurposing hit compounds or 1µM of DNMDP or 1µM anagrelide for 6h. Cells were collected and lysed with a modified radioimmunoprecipitation assay buffer (150 mM NaCl, 10% glycerol, 50 mM Tris·Cl pH 8.0, 50 mM MgCl2, 1% NP-40 detergent), supplemented protease and phosphatase inhibitors. Immunoprecipitation was performed using 2 mg of total protein lysates and 1 µg of anti-PDE3A antibody at 4°C overnight, followed by incubation with 7.5 µl each of Protein A and Protein G Dynabeads (10001D and 10003D; Thermo Fisher Scientific) at 4°C for 1h. Beads were washed with lysis buffer and proteins were eluted with 30 µl of lithium dodecyl sulfate–polyacrylamide gel electrophoresis loading buffer.

PDE3A and PDE3B enzyme activity assays. A commercial fluorescence polarization assay was performed using recombinant PDE3A and PDE3B by BPS Bioscience. Compounds were tested in duplicate at 9 concentrations in a half-logarithmic dilution series (top concentration of 10 µM) with a final DMSO concentration of 1%. The enzymatic reactions were conducted at room temperature for 30 min, 50 mM mixture containing co-crystallization buffer, 100 mM FAM-CAMP, a PDE enzyme and the test compound. After the enzymatic reaction, 100 µl of a binding solution (1:10 dilution of the binding agent with the binding agent diluent) was added to each mix and the reaction was performed at room temperature for 15 min. Fluorescence intensity was measured at an excitation of 485 nm and an emission of 528 nm using an Infinite M1000 Microplate reader (Tecan). Fluorescence intensity was converted to fluorescence polarization using the Magellan v6.1 software (Tecan).

Differential scanning fluorimetry (thermal shift) analysis of compounds binding to PDE3A. The gene for PDE3A (residues 677–1,141) was codon-optimized for Escherichia coli expression (GeneArt; Thermo Fisher Scientific) and cloned into an expression vector that attached an N-terminal polyhistidine sequence followed by a Tobacco Etch Virus protease cleavage site. The protein was expressed in E. coli and purified by affinity and size-exclusion chromatography. The polyhistidine sequence was removed by the Tobacco Etch Virus protease. PDE3A (5µM) was incubated for 20 min at room temperature with 100 µM of each compound. The reaction was performed using 20 µM HEPEs, pH 7.4, 150 mM NaCl, 500 µM TCEP, 5 mM MgCl₂, and 1% DMSO. After incubation, SYPRO orange (Thermo Fisher Scientific) was added to give a final concentration of 10x relative to stock concentration. Protein was tested using the LightCycler 480 (Roche Life Science). The temperature was increased from 25 to 95°C using a gradient of 0.06°C s⁻¹.

LS1034 CRISPR-Cas9 genome-wide knockout screen. The Brunello genome-scale sgRNA library was obtained from the Broad Institute GPP49. Virus was scale sgRNA library was obtained from the Broad Institute GPP49. Virus was

Tepoxalin competition assay. LS1034-Cas9–firefly luciferase cells were cocultured in a 1:1 ratio with LS1034–Renilla luciferase cells, both infected with ABCB1 sgRNA or a cutting control sgRNA. Cells were treated with 16µM tepoxalin (Wuxi) for 4 days and repeated with drug combination viability data. Tepoxalin luciferase activity was quantified using the Dual-Glo Luciferase Assay System (Promega Corporation) and measured using an EnVision plate reader. Data points were collected at day 0 and every 4 d until 12 d of treatment. The firefly to Renilla luminescence ratio was normalized to the initial day 0 measurement.

Tepoxalin cell permeability and stability assays. Compounds were incubated in medium alone or with 1 million cells ml⁻¹ at 37°C with gentle shaking for 3h. Following incubation, cell samples were centrifuged at 500g for 5 min and washed twice with cold PBS. Cells were resuspended in 130 µl of water. Samples were sonicated and centrifuged at 3,000g for 15 min at 20°C. 5µl of supernatant was combined with 45 µl of cell media, 50 µl of water and 50% of acetonitrile containing internal standard. Samples were centrifuged again and a final 100 µal aliquot was transferred to a 96-well plate for analysis. Samples were analyzed on an ultra-performance liquid chromatography–tandem mass spectrometry system consisting of an ACQUITY UPLC I-Class FTN (Waters) and Triple Quad 4500 System (SCIEX) with compounds detected by mode multiple reaction monitoring detection. Mobile phase A consisted of water with 0.1% formic acid (catalog no. 33015-1L; Honeywell), while mobile phase B consisted of acetonitrile with 0.1% formic acid. The gradient ran from 10 to 95% B over 0.8 min at a flow rate of 0.9 µl min⁻¹. An ACQUITY BEH C18 1.7 µm, 2.1 × 50 mm column was used, and the column was maintained at 65°C. Sample concentrations were determined using a standard curve and dilution quality-control samples prepared in a surrogate matrix. The Analyst v1.6.2 software was used for integration and calculation determination. For the stability study, 1 µM of tepoxalin was added to PBS, RPMI with 10% FCS or acetonitrile in duplicate and measured by mass spectrometry with time points prepared at 0, 24, 48, and 72h. The mass spectrometer was run in positive mode using multiple reaction monitoring detection for tepoxalin and the internal standard (75 nm midaomal).

Tepoxalin drug synergy/antagonism assays. Tepoxalin or paclitaxel were added in a dose–response matrix to LS1034 or REC1 cells, and viability was assessed by CellTiter-Glo. For the drug combination viability data, we first normalized CellTiter-Glo measurements by the median over DMSO wells on each plate. We then used the R package synergyfinder (version 1.8.0)20 to estimate Bliss synergy scores across all dose combinations, applying synergyfinder's default baseline correction method. Synergy values for each drug combination and cell line were summarized by the synergy score with the highest magnitude across dose combinations (maximum synergy). We verified that similar results were obtained qualitatively using other synergy models (for example, Loewe, highest single agent and zero interaction potency) and methods for aggregating synergy scores across dose combinations (for example, averaging).

ABCB1 activity assays. The ABCB1 antagonist assay was performed by Eurofins using published methods26. Briefly, MDR1-MDCK cells were incubated with test compounds and calcine acetoxyethyl. Change in calcine acetoxyethyl concentration was assessed by fluorescence measurement. A second assay, MDR1-MCF7 cell permeability, was performed by Cyprotex. Briefly, loperamide, a known MDR1 substrate, was added to the apical side of a cell monolayer and transport to the basal side was quantified over 60 min. Inhibition of MDR1-mediated transport of loperamide was assessed by adding tepoxalin or positive control (verapamil).

Transcriptional profiling by RNA-seq. LS1034 cells were seeded in 12-well plates. The following day, cells were treated in triplicate with 12µM tepoxalin or DMSO vehicle control for 6h. Wild-type SF295 cells and SF295 cells containing Cas9 and sgRNAs targeting GFP or MTF1 were plated in triplicate on a 6-well plate and incubated overnight. RNA was isolated using the RNeasy Mini Kit (QiAGEN) with DNase treatment. RNA quality was confirmed by Bioanalyzer (Agilent Technologies). Library preparation was performed by the Molecular Biology Core Facility at the Dana-Farber Cancer Institute using the KAPA RNA HyperPrep Kit (Roche). Nucleic acid was sequenced using an Illumina NextSeq 500 PE75 instrument. Gene-level expression values were obtained from RNA-seq using the TOPMed RNA-seq pipeline (version 1.1)21. RSEM (version 1.3.0) was used to generate transcripts per million gene-level expression quantifications. These tools were run using the FireCloud and Terra platforms22. Differential gene expression was calculated using the DESeq2 package (version 1.22.2)21.

Statistics and reproducibility. No statistical method was used to determine sample sizes. Compounds were plated for screening without regard to compound identity or ratio balanced. No data were excluded from the analysis, except individual data points flagged as assay failures by the process described earlier. Statistical tests are described in the text and figure legends with their associated sample sizes. PRISM data were generated without the investigators' knowledge of compound and cell line identities during screening. Investigators were not blind to compound and cell line identities during the analysis. Further
information on research design is available in the Nature Research Reporting Summary linked to this article.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

The PRISM Repurposing dataset, including screening data and all metadata, is available at the Cancer Dependency Map portal (https://depmap.org/repurposing). Raw and processed PRISM viability data are available from the Cancer Dependency Map portal (https://depmap.org/repurposing) and have been archived via figshare (https://doi.org/10.6084/m9.figshare.9393293). Interactive versions of Figs. 2a and 4 (with accompanying raw data) are also available on the Cancer Dependency Map portal; scatter plot source data are also deposited in figshare. The cell line features used for biomarker analysis are listed in Supplementary Table 10 and archived via figshare (https://doi.org/10.6084/m9.figshare.10277810). RNA-seq data have been deposited with the Gene Expression Omnibus (accession number GSE133299). All other data supporting the findings of this study are available from the corresponding author upon reasonable request.

**Code availability**

Data analysis was performed in R v.3.5.1 using custom-made or publicly available R packages. Individual packages are explicitly cited in the manuscript. The custom code is available upon request and from GitHub (https://github.com/broadinstitute/repurposing).

Received: 15 October 2019; Accepted: 6 December 2019; Published online: 20 January 2020

**References**


13. Meyers, R. M. et al. Computational correction of copy number effect and 4 (with accompanying raw data) are also available on the Cancer Dependency Map portal; scatter plot source data are also deposited in figshare. The cell line features used for biomarker analysis are listed in Supplementary Table 10 and archived via figshare (https://doi.org/10.6084/m9.figshare.10277810). RNA-seq data have been deposited with the Gene Expression Omnibus (accession number GSE133299). All other data supporting the findings of this study are available from the corresponding author upon reasonable request.


18. Analysis-Ready Standardized TCGA Data from Broad GDAC Firehose 2016_01_28 Run (Broad Institute TCGA Genome Data Analysis Center, 2016); https://doi.org/10.7930/C11G0K349.


Acknowledgements
We thank C. Yu, W. Hahn, B. Welpin, A. Bass, N. Gray, K. Stegmaier, E. Stover, T. Lewis, M. Mesleh, A. Burgin, S. Alper, G. Botta, M. Macaluso, P. Tsevetkov, X. Jin, K. Blakeslee, G. Ciolek and E. Lander for helpful scientific discussions. M. Passino, C. Zhu, K. Gore, M. Laird, C. Trapchis and E. Parikh generated the barcoded PRISM cell lines and performed the assays. K. Stumbraite assisted with STR fingerprinting. S. Johnson and J. Davis performed lysate processing and detection. S.E. Johnston and R. Singh provided analytical chemistry support. A. Vrcic, C. Sandland and S. Figueroa-Lazu assisted with compound management. G. Kugener and A. Gonzalez provided technical assistance. This study was supported in part by the Carlos Slim Foundation (Slim Initiative in Genomic Medicine for the Americas), the Next Generation Fund at the Broad Institute of MIT and Harvard (S.M.C.), the Conquer Cancer Foundation of ASCO Young Investigator Award (S.M.C.) and National Institutes of Health grants nos. U01 HG008699 (T.R.G and A.S.), U54 HL127366 (T.R.G and A.S.), KL2 TR002542 (S.M.C.) and K08 CA230220 (S.M.C.).

Author contributions

Competing interests
S.M.C., X.W., H.G., M.M., A.S. and T.R.G. receive research funding unrelated to this project from Bayer HealthCare. M.M. receives research funding from Ono and serves as a scientific advisory board and consultant for OrgiMed. M.M. has patents licensed to LabCorp and Bayer. M.M. and T.R.G. were formerly consultants and equity holders in Foundation Medicine, which was acquired by Roche. J.A.B. is an employee and shareholder of Vertex Pharmaceuticals. J.G.D. and A.T. consult for Tango Therapeutics. T.R.G. is a consultant to GlaxoSmithKline and is a founder of Sherlock Biosciences. Patent applications for the drug uses detailed in this manuscript have been filed. The other authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s43018-019-0018-6. Supplementary information is available for this paper at https://doi.org/10.1038/s43018-019-0018-6. Correspondence and requests for materials should be addressed to T.R.G. Reprints and permissions information is available at www.nature.com/reprints. Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. © The Author(s), under exclusive licence to Springer Nature America, Inc. 2020
Extended Data Fig. 1 | PRISM Repurposing assay and data processing overview. a, Lineage diversity of PRISM cell lines. The 489+ cancer cell lines tested span more than 23 tumor types. Lineages with fewer than 10 cell lines are listed on the right. b, Experimental protocol. Cell lines are grouped by doubling time into pools of approximately 25 cell lines. One pool is plated onto each assay plate. Compounds are transferred by pin transfer from a source compound plate (HTS and HTS002 screens), or cells are plated directly onto assay-ready plates generated by acoustic dispensing of compounds (MTS004, MTS005, and MTS006 screens). In either case, compound plates are shared by all replicates of each treatment condition. After incubation and lysis, all assay plates generated by a given compound plate are grouped and collapsed into 3 (HTS002, MTS005, and MTS006 screens) or 6 (HTS, MTS004 screens) detection plates so that each detection plate receives 1 or zero copies of each pool. Ten control barcodes are then spiked into each detection plate well (HTS002, MTS005, and MTS006 screens). Detection plates are amplified by PCR and detected using Luminex FLEXMAP 3D instruments. c, Data processing workflow. Median Fluorescence Intensity (MFI) values are calculated from fluorescence values for each replicate-condition-cell line combination and are log2-transformed. Assay plates wells are normalized, median-collapsed, and compared to the normalized medians of other assay plate wells in the same well position that have been dosed by the same compound plate. A robust z-score is calculated, and assay plate wells with a |z-score| > 5 are filtered. Strictly standardized mean differences (SSMD) are calculated between positive and negative control conditions for each cell line on each assay plate. Cell line-assay plate combinations with SSMD < 2 are filtered by a control-separation filter to generate the log MFI data matrix. In datasets with control barcodes added, data are normalized with respect to the median of control barcodes to generate the MFI normalized data matrix. Data are DMSO-normalized and pooling artifacts are corrected using ComBat to generate the log fold change data matrix. Up to 3 independently treated plates (range 1-3 based QC filtering) in one screen are median-collapsed to generate the collapsed log fold change data matrix.
Extended Data Fig. 2 | Outlier pool QC filter to detect pool-level failures. **a**, Primary screen QC pass rate by pool. The fraction of treated assay plate-wells that pass the outlier filter is indicated. Cell line log MFI data are median-centered, and the medians of assay plate-wells are compared within each well-detection plate combination. Extreme outliers with |robust z-score| > 5 are filtered. **b**, Primary screen QC pass rate by detection plate. **c**, Primary screen QC pass rate of assay plates. Overall pass rate was high (median 98.6%, minimum 83.8%). **d**, Secondary screen QC pass rate by pool. **e**, Secondary screen QC pass rate by detection plate. The pass rate is above 95% for 81% of detection plates. **f**, Secondary screen QC pass rate by plate-pool combinations. 3 replicate plates are combined for visualization. Overall pass rate was high (median 99.1%, mean 94.3%), where failures are almost exclusively coming from 7 detection plates, implying failure at final detection step. Across the screen, 5.7% of the pools are filtered out as outliers.
Extended Data Fig. 3 | Control-separation QC filter to detect cell line failures. 

a, Primary screen QC pass rate by cell line. SSMD of log MFI values is calculated between positive control and negative control treatments for each cell line on each plate. Data from cell line-plates with SSMD < 2 are filtered. 
b, Primary screen QC pass rate by detection plate. 
c, Primary screen QC pass rate of assay plates. Overall pass rate was high (median 99.2%, mean 94.3%). Some detection plates show higher failure rates for specific detection pools, implying failure at final detection step. The bulk of the filtered data was from two detection plate-detection pools (PREP013_X2 and PREP003_X1 in detection pool 8).

d, Secondary screen QC pass rate by cell line. SSMD of log MFI values is calculated between positive control and negative control treatments for each cell line on each assay plate. Data from cell line-plates with SSMD < 2 are filtered.

e, Secondary screen QC pass rate by detection plate.

f, Secondary screen QC pass rate of assay plates. Overall pass rate was lower than primary (median 99.88%, mean 79.1%). Similar to the primary screen, the main mode of failure is platewise failures.
Extended Data Fig. 4 | Number of well replicates passing QC in the PRISM screens. **a**, Number of individual cell assay well replicates (max n = 3) that pass QC filters in the primary screen, grouped by compound plate and cell line quality. 86% (497 out of 578) of the cell lines have at least one passing replicate for all compound plates. Identity of cell lines with lower quality data are listed at the bottom. **b**, Number of individual cell assay well replicates (max n = 3) that pass QC filters in the secondary screen. 95% (463 out of 489) of cell lines have at least 1 passing replicate on at least 85% (30 out of 35) of the compound plates.
Extended Data Fig. 5 | Comparison of PRISM viability data to reference datasets. a, Pairwise Pearson correlations between drug response AUCs of publicly available datasets. AUC values were recomputed for 84 compounds and 318 cell lines (median 236 cell lines per compound) over the same dose-range for each compound-cell line pair in all 3 datasets (GDSC, CTD2, REP), and capped at 1. The Pearson correlation across shared compound-cell line pairs of the datasets is above 0.6. 44.8% of cell line-compound pairs show inactivity (AUC > 0.8) in all three datasets (n = 16650 compound-cell line pairs). b, Pairwise Pearson correlations between drug response AUCs after removing inactive cell line-compound pairs. Pearson correlations were re-calculated after filtering out the inactive data points (AUC > 0.8 in all three datasets) (n = 9188 compound-cell line pairs). c, Compound-wise correlation between publicly available datasets. Correlation between PRISM data and other datasets is similar to correlation between other datasets. Points represent Pearson correlations and error bars represent 95% confidence intervals computed using Fisher’s z-transform. GDSC vs. CTD2 is in blue and REP vs. GDSC/CTD2 is in red. The number of cell lines shared by all three datasets is shown after each drug name. Paired t-tests on compound-wise correlations show statistically significant (two-sided p-value: 0.012 and 0.014 for top and bottom, respectively) but small mean of differences (0.049 and 0.039 for top and bottom, respectively). The number of data points used to compute each correlation is given in the figure for each compound.
Extended Data Fig. 6 | PRISM Repurposing noise quantification. a, Cell line standard error estimates across vehicle-treated wells on PRISM plates treated with DMSO (n = 489 cell lines). Log fold change standard errors are estimated for each cell line using DMSO-only plates included in the MTS006 screen (n = 384 × 3 replicate wells for each cell line). b, Comparison of standard error estimates across screens. The error estimate calculation is repeated for each screen using DMSO wells on standard compound plates (n = 32 replicate wells per plate), except for MTS006, which uses DMSO-only plates (n = 384 replicate wells per plate). Higher noise levels are observed in the initial high-throughput screens HTS001 (n = 578 cell lines) and HTS002 (n = 489 cell lines) compared with the medium-throughput screens MTS004 (n = 578 cell lines), MTS005 (n = 489 cell lines), and MTS006 (n = 489 cell lines). Upper box limits, center lines, and lower box limits correspond to 75th, 50th, and 25th percentiles, respectively. Whiskers extend from the box limits to the most extreme value up to 1.5 IQR from the median. All cell lines are depicted as points, regardless of outlier status. c, Comparison of estimated standard error of vehicle control wells between high-throughput pharmacogenomic datasets. The average standard error across cell lines (n = 301 cell lines for PRISM versus CTD2 and n = 197 for PRISM versus GDSC) is indicated by dashed lines with standard deviation in parentheses. d, Relationship between drug selectivity and replicate reproducibility in PRISM. Average Pearson correlation between replicates for each compound, dose, and screen combination is stratified by mean bimodality coefficient. For comparison, the null distribution for randomly paired compounds is shown in blue.
Extended Data Fig. 7 | Generation of knockout cell lines by CRISPR/Cas9 editing. 

a, SF295 cells were transduced with multiple guides targeting the MTF-1 gene. Following selection, genomic DNA was isolated, and the targeted region was amplified by PCR. Results from the NGS CRISPR assay are shown as percent indel formation.
b, Differentially expressed genes in SF295 glioma cells following MTF-1 knockout by CRISPR/Cas9. Loss of MT1E, MT1X, and MT2A expression was observed upon MTF-1 knockout. Gene expression across three independent cell wells per cell line were measured by mRNA sequencing. Two-sided p-values for differential gene expression following MTF-1 knockout vs. parental cell line were calculated with DESeq2 and corrected for multiple hypothesis testing using the Benjamini-Hochberg method.
c, Drug sensitivity of SF295 cells with and without MTF-1 knockout. MTF-1 does not alter sensitivity to control chemotherapeutic bortezomib. Mean viability across 3 independently treated wells is shown, with standard deviation indicated by error bars.
d, Western immunoblot validation of SLC26A2 knockout in OVISE ovarian and A2058 melanoma cancer cell lines. The SLC26A2 protein is known to migrate across a range of molecular weights due to glycosylation. Results are representative of two independent experiments.
e, OVISE cells were transduced with multiple guides targeting the SLC26A2 gene. Indel frequency at the SLC26A2 CRISPR Cas9 cut sites was assessed by NGS CRISPR assay.
f, ABCB1 western blot with and without CRISPR knockout of ABCB1 in the LS1034 colon cancer cell line. Western blot was performed once.
g, Percent indel formation at genomic cut site in LS1034 ABCB1 CRISPR knockout lines assessed by NGS CRISPR assay.
h, Cellular viability of LS1034 CRISPR knockout lines after treatment with tepoxalin for 5 days. Mean viability across 3 independently treated wells is shown, with standard deviation indicated by error bars.
**Extended Data Fig. 8 | Tepoxalin mechanistic studies.**

**a**, ABCB1 western blot with and without overexpression of ABCB1 in the Kuramochi ovarian cancer cell line. Western blot was performed once with three independent samples. **b**, Cellular viability of Kuramochi wild type and ABCB1-overexpressing cells after treatment with tepoxalin for 8 days. Mean viability across 3 independently treated wells is shown, with standard deviation indicated by error bars. **c**, (Left) Single-agent dose-response curves for LS1034 cells treated with tepoxalin and zosuquidar for 5 days. Two replicates were averaged. (Middle) Dose response curves for tepoxalin in combination with varying doses of zosuquidar (indicated by different colors) for 5 days. Data are shown for tepoxalin doses above 470 nM (the range indicated by the vertical dashed lines above). (Right) Bliss synergy scores estimated for each dose combination, showing strong antagonism by zosuquidar at tepoxalin doses above ~5 μM. Two combinations were not tested (NT). **d**, Maximum synergy score is shown across several drug combinations measured in both LS1034 and REC1 cell lines. Both tariquidar and elacridar were strongly synergistic in combination with paclitaxel, while all MDRI inhibitors tested were antagonistic in combination with tepoxalin. **e**, Cellular permeability study of tepoxalin and RWJ20142 in LS1034 colon cancer cell lines with and without ABCB1 knockout. Each indicated cell line or media-only control was treated with 20 μM tepoxalin or RWJ20142 for 3 hours. Tepoxalin and RWJ20142 concentrations in cell lysate or medium were determined by liquid chromatography-mass spectrometry. Mean viability across 3 independently treated wells is shown, with standard deviation indicated by error bars. Shaded dots indicate underlying data. **f**, Stability study of tepoxalin in three different vehicles over 72 hours. Percent of tepoxalin remaining is indicated at each timepoint. **g**, ABCB1 antagonism assay using calcein AM fluorescence. MDCKII cells were treated with indicated concentration of tepoxalin or RWJ20142. Mean percent ABCB1 inhibition across 3 replicates is shown. **h**, ABCB1 transport-based activity assay. Basolateral transport of the ABCB1 substrate loperamide was assessed using a monolayer of MDCK-ABCB1 cells in the presence of tepoxalin or RWJ20142. Mean percent ABCB1 inhibition across 2 replicates is shown.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  - Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Luminex xPONENT version 4.2 was used to collect data.

Data analysis

Data analysis was performed in R version 3.5.1, using custom or publicly-available R packages. Individual packages are explicitly cited in the manuscript and listed below. The custom code is available upon request and from GitHub (https://github.com/broadinstitute/repurposing). Graphpad PRISM version 8 was used for curve fitting and dose response plots.

The following software packages were used for data analysis:

- DESeq2 v3.10
- Limma v3.38.3
- Atlantis [cancerdatasci/atlantis SHA: 167fc0a338fc2686a27463eb555466fe316d91b0 on github]
- MAGeCK v0.5
- svge 3.30.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw and processed PRISM viability data are available from the Cancer Dependency Map portal (https://depmap.org/repurposing) and archived via Figshare
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-list.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size
- No sample size calculation was performed. PRISM cell lines were selected based on ability to be transduced with lentivirus and adapt to shared culture conditions. In the primary screen, data was collected for 578 cell lines, 489 cell lines were included in the secondary screen, based in part, on primary screen QC performance. The 578 cell line panel spans a range of multiple cancer types, patient age (pediatric and adult cancer lines), and gender.

Data exclusions
- A standardized QC process was employed for the processing of PRISM data. The rationale for filtering data was to remove low quality data points due to technical failures. Replica-level outliers were filtered using the approach fully described in Methods and Extended Data Fig. 2-4. All filtered data is still included in the raw data files.

Replication
- The PRISM screen was performed with 3 independently treated well replicates. Multiple independent cell treatments were performed for conformational studies as indicated in the Methods. Results shown are indicative of multiple independent experiments as indicated in the figure legends.

Randomization
- All cell lines were tested with active treatments and a vehicle control. No individual experimental groups were used that would necessitate a randomized design.

Blinding
- PRISM data was generated without the investigators' knowledge of compound and cell line identities during screening. Investigators were not blind to compound and cell line identities during analysis in order to interpret results within the context of known drug activities.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- n/a
- Antibodies
- Eukaryotic cell lines
- Palaeontology
- Animals and other organisms
- Human research participants
- Clinical data

Methods

- n/a
- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging

Antibodies

- The following antibodies were used: polyclonal rabbit anti-PDE3A from Bethyl Laboratory (A302-740A, 1:1000 dilution), monoclonal mouse anti-V5 from Life Technologies (R660-25, 1:5000 dilution), monoclonal rabbit anti-ABC1 (D3H1C) from Cell Signaling Technology (1268), 1:1000 dilution), monoclonal mouse anti-β-Actin (B100010) from Cell Signaling Technology (3700, 1:1000 dilution), and monoclonal rabbit anti-SLC26A2 Antibody (3F6) from Novus Biologicals (H00001836-M04, 1:1000 dilution).

Validation

- We validated the ABC1 and SLC26A2 antibodies by performing gene knockout by CRISPR/Cas9 as reported in this manuscript. For the anti-V5 antibody, the manufacturer provides western blot validation by expressing a V5-tagged transgene. For the PDE3A antibody, IP validation is available from the manufacturer and in the literature (PMID26656089). The actin antibody has validation western blot data available from the manufacturer and is widely cited in the literature.
Eukaryotic cell lines

Policy information about cell lines

Parental cell lines were obtained from the Broad-Novartis Cancer Cell Line Encyclopedia (CCLE) project prior to PRISM barcoding (see https://portals.broadinstitute.org/ccle for original sources). For follow-up studies, LS1034, HeLa, HEK293T cells were purchased from the American Type Culture Collection. REC1, Sf295, OVISE, COLO320, BEN, HT29 and SNU449 were provided by CCLE. Wildtype and ABCB1 overexpressing (piX_317 vector) Kuramochi cell lines were gifts from Elizabeth Stover (PMID31462500). HeLa PDE3A CRISPR KO cells (PDE3A/- cells) were previously described (PMID26656089). LS1034, Sf295, A2058, COLO320, SNU449 and OVISE cell lines derived with Cas9 were provided by the Broad Cancer Dependency Map.

Authentication

STR fingerprinting was performed by Genetica using the PowerPlex® 16 HS system (Promega). STR profiles were compared with STR profiles reported by vendors and in literature. Misidentified cell lines or other STR conflicts are listed in Supplementary Table 9. These cell lines are flagged in the data files and are not shown by default in the website interface.

Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination.

Commonly misidentified lines (See ICLAC register)

KPL-1/MCF-7 and NCI-2077/NCI-H1581 have been listed in ICLAC as misidentified or derivative lines.