# Scalable and universal prediction of cellular phenotypes

Yuge Ji[1,2]*, Alejandro Tejada-Lapuerta[1,3]*, Niklas A. Schmacke[1], Zihe Zheng[1], Xinyue Zhang[3], Simrah Khan[1,4], Ina Rothenaigner[5], Juliane Tschuck[5], Kamyar Hadian[5], Fabian J. Theis[1,2,3*]

[1]Institute of Computational Biology, Helmholtz Munich, Germany
[2]School of Life Sciences, Technical University of Munich, Munich, Germany
[3]School of Computing, Information and Technology, Technical University of Munich, Munich, Germany
[4]Ludwig Maximilians University, Faculty of Biology
[5]Research Unit Signaling and Translation, Helmholtz Munich, Neuherberg, Germany

*Correspondence to fabian.theis@helmholtz-muenchen.de
*equal contribution

Biological systems can be understood by perturbing individual components and studying the system's response. Cell biology experiments are defined by the applied treatment, cellular state, and the assayed phenotype. Given the vast number of possible combinations, testing every scenario is impractical. We present Prophet, a transformer-based computational model for cellular phenotype prediction. Prophet learns a representation of the cell biology experiment space, enabling it to predict the outcomes of untested small molecule or genetic perturbations in new cellular contexts across diverse phenotypes including gene expression, cell viability, and cell morphology. Its scalable architecture facilitates training across independent assays, using transfer learning to enhance performance across phenotypes. In vitro validation shows Prophet's potential to guide experimental design, making it a valuable tool for accelerating biological discovery.

## Introduction

High-throughput screening and multiplexed phenotypic readouts have revolutionized biological experimentation. They enable simultaneous testing of thousands of compounds across numerous biological contexts, accelerating the pace of discovery. For instance, small molecule and CRISPR screens have been conducted to systematically interrogate the biological mechanisms of cancer resistance (*1*, *2*), identify new drug targets (*3*, *4*), capture synergistic interactions (*5*), and deconvolve drug mechanisms of action (*6*–*9*). While statistical and AI-based analyses have facilitated discoveries in each screen, a key limitation of past models is the lack of integration across screens (*10*, *11*), despite considerable overlap in experiments—for instance, the same cell line treated with different compounds or genetic perturbations, or the same experiment measured with different readouts. Learning across all of these screens could

provide notable increases in the efficiency of identifying effective and cell-context specific compounds while simultaneously making findings more robust across technical variation (*12*).

Given the immense combinatorial space inherent to biological experimentation, accurate *in silico* modeling is needed to guide the selection of relevant future experiments and reduce the number of early-stage experiments needed to yield new biological insights(*13*). While previous methods offer tailored solutions for specific experimental contexts (*11*, *14–18*), we propose a general approach to treatment response prediction which can fully leverage the vast data resources now available. Inspired by the success of transformers in solving complex problems across various domains by modeling data as a set of tokens to be transformed towards an outcome (*19–21*), we develop a novel approach to cellular response prediction: modeling the experimental space using set-based representations. We universally decompose each experiment into a unique combination of three fundamental elements—the cellular state, the treatments being performed, and the intended phenotypic readout. This approach forms the basis for our model, Prophet (**Pr**edictor **o**f **phe**not**y**pes), a transformer-based regression model that learns the relationships between these factors. Prophet's framework enables it to be pretrained on 4.7 million experiments across a broad spectrum of phenotypes from multiple independent datasets, producing accurate predictions that empower downstream applications, such as nominating additional chemotherapeutic drugs for *BRAF*-mutant melanoma and reducing the number of experiments needed for viability screens by at least 60x. Together, its scaling behavior and strong predictive capacity establish Prophet as a robust general model for treatment response in biological systems.

# Results

## A unified framework for experimental cell biology

We consider the outcome of a cell biology experiment as the composite of three essential components: the cell state the experiment is performed in, the applied treatment, and the phenotypic readout used to measure the effect of the perturbation (Fig. 1a, b). The cellular state represents the basal state before treatment. For instance, most drug screens are performed on cell lines that we characterize as the cell state. The treatment is the specific intervention or perturbation applied to the cells during the experiment. Commonly, these are genetic perturbations (e.g. CRISPR/Cas9 knockouts, CRISPRi/a) or drug perturbations with small molecules. Finally, the phenotypic readout is the measurable outcome used to assess the effect of the applied treatment on the cell state; for example a morphological measurement such as cell size, or the $IC_{50}$ of a compound. To model the relationship between these elements in a generalizable way, we developed Prophet, a supervised transformer-based deep learning model (Fig. 1b, c). To train it, we collected 9 perturbational datasets to create the largest compendium of publicly available screening datasets to date: SCORE(*3*, *22*), LINCS(*23*), JUMP(*8*), GDSC(*3*, *22*, *24*), GDSC$^{2(25)}$, CTRP(*26*), PRISM(*1*), Horlbeck et al.(*27*) and Saunders et al.(*28*) (fig. S1, Table S1). For example, we consider one measurement of the $IC_{50}$ of the drug isoniazid in HeLa cells a single experiment that is passed to Prophet as one datapoint. Measuring cell morphology after the same treatment in the same cell line would be a different data point, and so would

measuring the $IC_{50}$ of the same drug in another cell line (Fig. 1c). We hypothesized that when trained across a sufficient amount of experimental data, a single model could be employed to make predictions for any compound or genetic treatment, and for any cell line or cell type (Fig. 1).

Prophet's architecture consists of 8 transformer encoder units (*29*, *30*) with 8 attention heads per layer and a feed-forward network with a hidden dimensionality of 1,024 to generate a 512-dimensional embedding of each experiment (Fig. 1c, fig. S2). Prophet leverages existing knowledge of cellular states and treatments by projecting prior knowledge-based representations into a common token space using neural networks as tokenizers. The readout representations are modeled as learnable embeddings, directly projected in the token space. Through self-attention, Prophet effectively captures relationships between treatments, cellular states, and phenotypic outcomes, facilitating nuanced learning and prediction. The output of the transformer is fed into a feed-forward network that predicts the phenotype of interest. Prophet is trained end-to-end using supervised learning to minimize mean square error (MSE) between the predicted and measured phenotypes.

To make prior knowledge available for training Prophet, we create representation vectors for each axis of the experimental space, encoding the cellular environment the experiment is conducted in, and the treatments that are applied, to serve as input to the model.
To represent cell state, as most small molecule and CRISPR screens are performed in cell lines, we use bulk RNA sequencing data from the Cancer Cell Line Encyclopedia (CCLE) (*31*), which covers 1,406 cell lines. For *in vivo* treatments, we assume each cell type in a tissue represents one cell state and encode them using their RNA expression profile after pseudobulk aggregation.
For treatments, we use two different methods of generating embeddings depending on whether the treatment is a small molecule or genetic perturbation. For small molecules, we create descriptive embeddings by concatenating their RDKit2D and Morgan fingerprints (*32*, *33*). For genetic perturbations, we build a vector that represents the target gene as the concatenation of a vector capturing its transcriptomic variability and a vector that represents its genomic sequence (Methods). We obtain the transcriptomic variability vector by transposing the CCLE "cell line x gene expression" matrix and finding a low dimensional representation of the genes using PCA. We derive the genomic sequence vector from Enformer's embeddings (*34*) for the entire coding sequence of the gene.
Since there is no prior knowledge suitable to represent an embedding of most phenotypes, we model them as a learnable embedding through one hot encodings and a linear layer. However, if there exists a suitable prior knowledge-based representation of the phenotype, it can be used as input to a tokenizer network, similar to treatment and cell state representations.

Prophet is capable of jointly modeling diverse cellular phenotypes used as readouts in experiments. We train Prophet to predict cell viability, compound $IC_{50}$, Cell Painting morphology features (*35*), mRNA transcript abundance, and cell type proportion. Its transformer-based architecture allows Prophet to handle an arbitrary number of perturbations, from singletons to any combination, so that datasets with single treatments and combinations can be modeled

under the same learning framework. Additionally, utilizing prior knowledge-based embeddings enables Prophet to predict *in silico* outcomes for experiments performed in unseen conditions, extrapolating to unseen perturbations or cellular states.

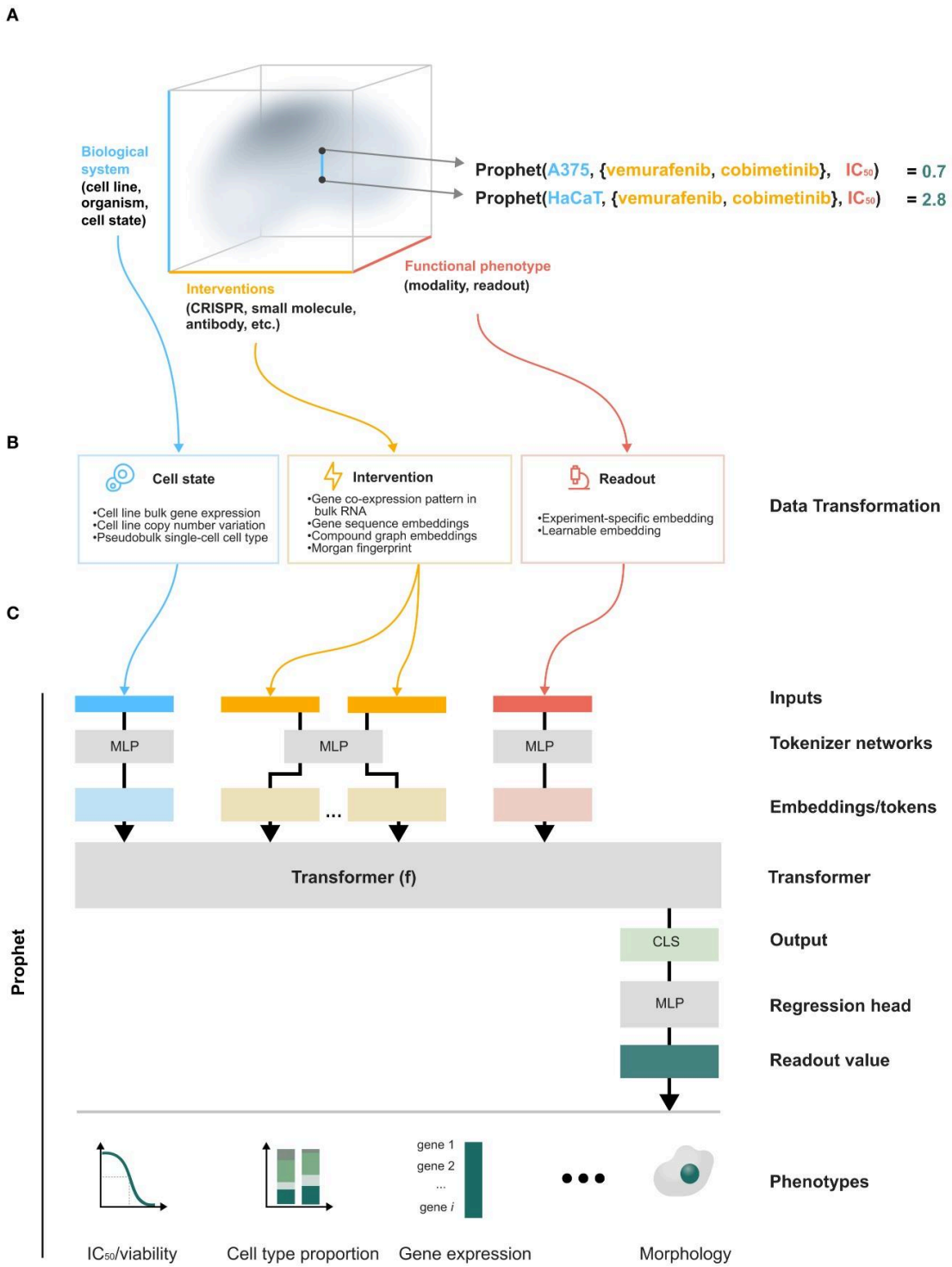


**Fig. 1. Prophet is a universal phenotype predictor.** (**A**) Prophet models the experimental space along

three axes, namely cellular state, treatments, and phenotypic readout. Each point in the experimental space is mapped to its corresponding phenotypic readout. We illustrate an example of how Prophet samples two different points of the experimental space. The two points differ in the biological system these experiments were performed in: HaCaT and A375 cells. This flexibility enables Prophet to be trained across datasets. (**B**) Cellular states and treatments are passed to Prophet as prior knowledge-based representations. Phenotypes can be either modeled as a learnable embedding or grounded on prior knowledge. (**C**) Prophet relies on a transformer-based architecture that leverages self-attention and permutation equivariance to represent the experimental context as a set and obtain a meaningful representation that maps to the phenotypic readout. It consists of tokenizer networks that take prior knowledge to create tokens, the transformer to which those tokens are fed and a small regression head that uses the output of the transformer to get a readout prediction. Prophet can predict any quantitative phenotype, including cell viability, gene expression and cell morphology.

## Transfer learning improves *in silico* experimental prediction

We assessed Prophet's efficacy in predicting the outcome of unseen experiments. We conducted five-fold cross-validation with three validation seeds (Methods, fig. S5a) and compared the average coefficient of determination (R2) and the Spearman correlation of Prophet against three baseline architectures: random forest, a multilayer perceptron, and linear regression. For each method we report the difference in R2 and Spearman correlation to the mean estimator (Methods). The mean estimator serves as a global approximation of a treatment or cell state response based on prior experiments (fig. S5a, Methods) and is considered a robust baseline in drug response prediction (*11*).

We first trained specialized models on each dataset individually, with the same architecture but a smaller number of parameters to prevent overfitting (Fig. 2a, Methods). We call these models “Prophet-individual” (for hyperparameter tuning, see Supplementary Methods, Table S2). These Prophet-individual models generally outperform baseline methods across the board (Fig. 2b). Then, to benchmark Prophet’s transfer learning capabilities, we conducted experiments training jointly on two datasets (Prophet-two-dataset). We combined training on $GDSC^2$[25], a small molecule combinatorial perturbation screen reporting $IC_{50}$ values, with PRISM(*1*), a much larger single-treatment screen relying on a CellTiterGlo viability readout. PRISM contains 8 of the 61 compounds and 76 of the 110 cell lines in $GDSC^2$. Performance for both unseen treatments and unseen cell lines in $GDSC^2$ improved by 3% and 28%, respectively, after joint training with PRISM. These results suggested Prophet could benefit from joint training on additional datasets and prompted us to increase the amount of datasets used during pretraining.

We then pretrained our largest Prophet model on 4.7 million experimental conditions, encompassing 26k treatments, 1,550 cell states, and three different readouts ($IC_{50}$, viability and RNA expression) across seven small molecule and genetic perturbation screens (fig. S2a, b, Table S1). Afterwards, we fine-tuned the model for up to 20,000 steps with a batch size of up to 256 observations and a low learning rate (Fig. 2a, Methods, Table S3). This approach follows the same strategy as current foundation models (*19*, *36*, *37*), which are pre-trained on large amounts of data and then fine-tuned for specific datasets using the pre-trained model as a backbone.

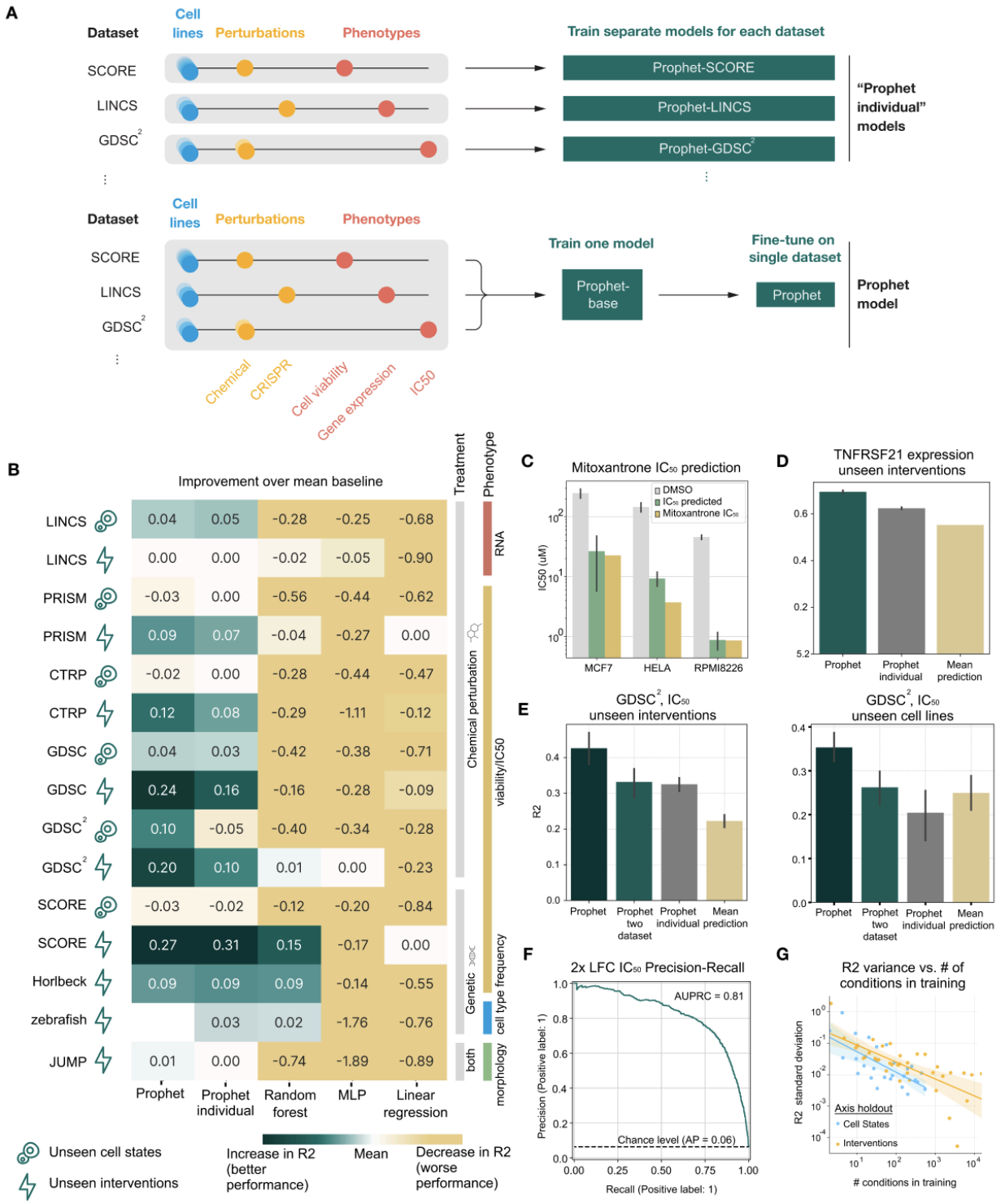


**Fig. 2. Prophet's performance and scaling across diverse datasets.** (**A**) Prophet-individual models (top) are Prophet models trained only on a single dataset. On the other hand, the full Prophet model (bottom) is pretrained using all datasets and then fine tuned on the dataset of interest. (**B**) Improvement in the R2 score over the mean baseline of the models evaluated: linear regression, multilayer perceptron (MLP), Random Forest, single-assay Prophet (Prophet-individual) and fine-tuned Prophet (Prophet). (**C**) Prophet accurately predicts the $IC_{50}$ value of the unseen compound mitoxantrone, a clinically approved

compound, against different cancers. (**D**) The fluorescent RNA readout of the gene TNFRSF21 is not part of the pre-training corpus. However, fine-tuning Prophet on it still yields better results than the mean estimator and than training a Prophet-individual model only on this dataset. (**E**) In $GDSC^2$, fine-tuned Prophet outperforms Prophet-individual, the mean estimator and also a Prophet-two-dataset model trained jointly in $GDSC^2$ and PRISM. Notably, Prophet is fine-tuned for just 20,000 steps with a batch size of 256 observations – much fewer optimization steps than the other models, which highlights the effectiveness of the pre-training stage. (**F**) The area-under the precision-recall curve (AUPRC) is 0.81 for *hit* identification in GDSC - defining each *hit* as a compound that causes at least a 2x fold change in cell viability. (**G**) The variance in the R2 score obtained by Prophet is reduced as more parts of the experimental space - more cell states and more treatments - are sampled during training.

We find that pretraining Prophet on all datasets and fine tuning for the dataset of interest consistently matches or outperforms baseline models, including Prophet-individual models, across all datasets, with baseline models frequently failing to outperform the mean estimator (Fig. 2b). Prophet demonstrates notable performance not only in predicting responses to small molecule monotherapies (GDSC, CTRP, PRISM) but also in predicting the outcomes of unseen genetic perturbations (SCORE). In addition, Prophet successfully predicts the effect of the combination of two drugs and the combination of two CRISPR perturbations (GDSC combinations and Horlbeck, respectively). To showcase Prophet's ability to extrapolate to unseen scenarios we predict the $IC_{50}$ of Mitoxantrone, a clinically approved compound indicated for multiple cancers (*38*, *39*), which does not appear in any dataset in training. Still, Prophet's de novo predicted $IC_{50}$ for Mitoxantrone reflects its efficacy in different cancer types (Fig. 2c), and is corroborated by the $IC_{50}$ values measured in GDSC (*24*) as well as other publications (*40*, *41*). Furthermore, in addition to viability, Prophet accurately predicts diverse and complex measurements including cell morphology (JUMP) and transcriptomics (LINCS) (Fig. 2b). This generalizability serves as evidence of Prophet's ability to enable transfer learning, leveraging information across multiple screens. As an example, pretraining Prophet and fine-tuning it on $GDSC^2$ outperforms Prophet-individual by 70% in predicting unseen treatments and by 30% in predicting experiments on unseen cell states. Most notably, it also outperforms a Prophet model trained on PRISM and $GDSC^2$ (Prophet-two-datasets) by 28% in predicting unseen treatments and by 29% in predicting experiments on unseen cell states (Fig. 2e).

Prophet's transfer learning capability is not limited to phenotypes seen during the pre-training stage. We did not pre-train Prophet on any morphological measurement, but Prophet fine-tuned on JUMP outperformed both the Prophet-individual model trained only on JUMP and the baseline (Fig. 2b). We observed the same behavior with unseen RNA readouts: the mRNA abundance readout of the gene TNFRSF21 from LINCS was not included in the pre-training corpus; nonetheless, the fine-tuned version of Prophet is 25% better than the baseline and 10% better than a single-dataset Prophet model trained uniquely on TNFRSF21 readouts (Fig. 2d).

The prediction of the most likely treatments to be a *hit,* defined as an intervention that results in a desired biological effect, can reduce the number of total screening experiments needed. Critically, Prophet's performance is not limited to accurate prediction of values close to the mean but Prophet also accurately predicts extreme values—hence, possible *hits* (fig. S5b). Therefore,

we can use Prophet to generate *in silico* outcomes for treatment or cell state combinations not yet tested, and use these values to suggest experiments which are most likely to produce a desired phenotype. As a case example, we take Prophet's *in silico* simulated viability values and compute the increase in the number of *hits* over randomly selecting compounds from the GDSC library. We define a *hit* as a compound that induces greater than 2x fold-change in viability as compared to DMSO in a given cell line (Supplementary Methods). Using Prophet to predict viability for drugs in unseen cell lines in GDSC shows an improvement in R2 of .04 over the mean estimator (Fig. 2b). However, this increase translates into a 13x increase in the number of hits when performing experiments suggested by Prophet. Here, Prophet achieves an area-under the precision-recall curve (AUPRC) of .81 (Fig. 2f). When identifying hits with a more stringent viability fold-change of at least 4x, the potential of Prophet is even more apparent: here, Prophet leads to a reduction in performed experiments of 60x compared to identifying hits from randomly selected compounds (Supplementary Methods). Prophet was conservative in its predictions and had no false positives among its top-ranked 1% of hits. In addition to precision-recall, we further evaluate using Spearman rank (fig. S5c). Prophet outperforms the baseline models and the mean in all instances in which it also did the same using R2. This demonstrates that Prophet not only accurately predicts phenotypic values but also effectively captures the underlying ranking of responses, indicating its ability to make reliable comparisons across different experimental conditions.

To investigate how the uncertainty of Prophet varies with the number of conditions (cell states and treatments) seen during training, we trained an ensemble of Prophet models. We found a clear trend: the more treatments and cell states seen by Prophet, the higher the confidence in the predictions (Fig. 2g). Likewise, we also studied how the performance of Prophet changes as a function of the number of conditions seen during training: we trained Prophet-individual models to predict the effect of either unseen treatments or unseen cell states, including 5-80% of the data along the respective axis. The amount of data required to outperform the mean estimation varies per dataset, but as expected we observe that Prophet's predictions improve with additional training data (fig. S4, S5).

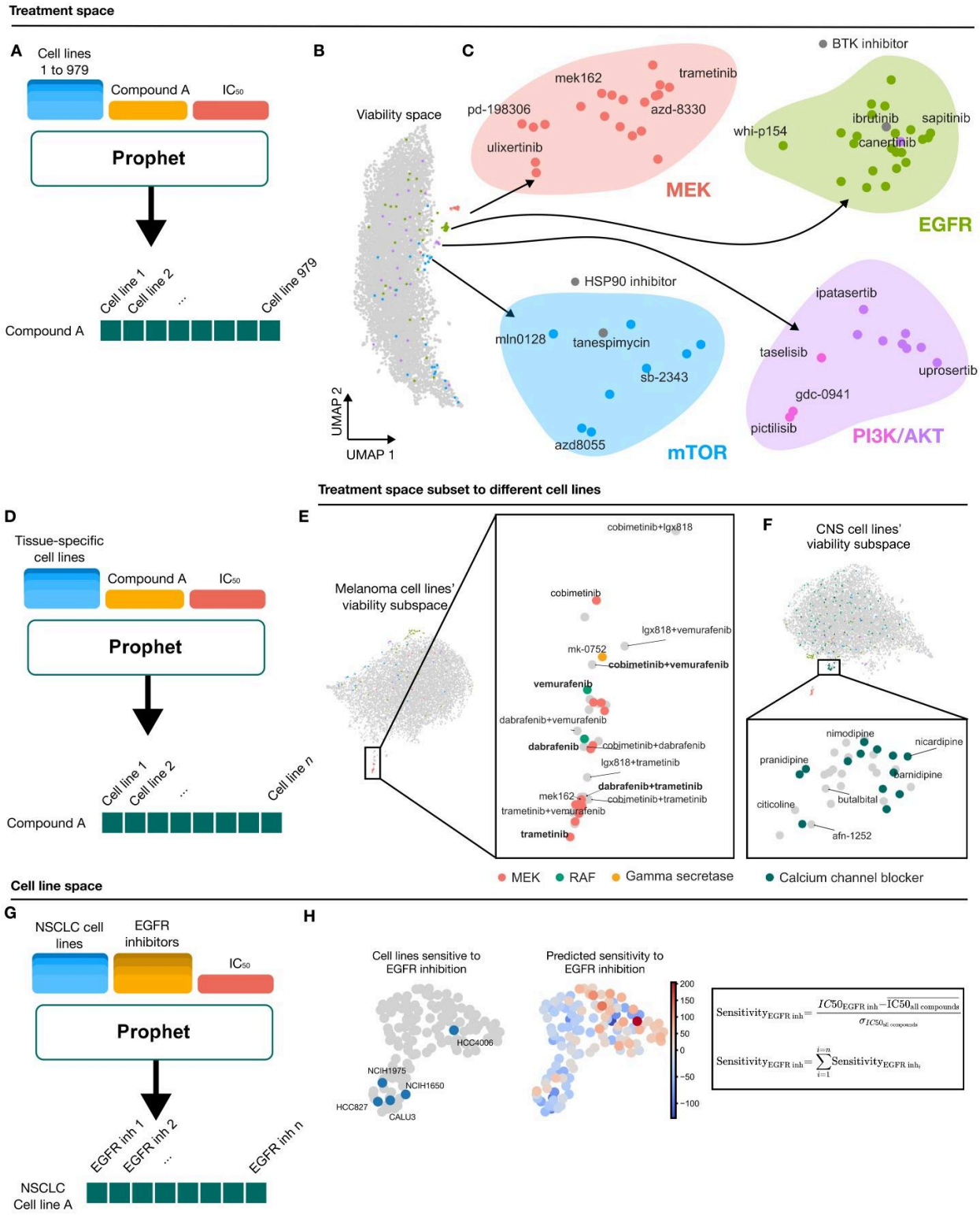


**Fig. 3. Prophet learns a map of the experimental space.** (**A**) Schematic of how Prophet generates a complete experimental space *in silico*. For the treatment space, Prophet predicts the $IC_{50}$ for each compound or compound combination across 979 cell lines. Each single or combinatorial treatment is then

represented as a 979-dimensional vector, in which each dimension is Prophet's $IC_{50}$ prediction for that compound in one cell line. We call this treatment representation an "activity-based fingerprint". (**B**) Two dimensional visualization (UMAP) of activity-based fingerprints of 8,620 small molecule treatments, each represented as one point. (**C**) Compounds with differential effects can cluster into four different biological signaling pathways involved in cell proliferation: MAPK, EGFR, PI3K/AKT and mTOR (Table S6). (**D**) Prophet can create experimental subspaces by constraining the cell lines or compounds. For the melanoma and brain cancer subspaces, only cell lines representing the respective tissues are included. (**E**) $IC_{50}$ treatment subspace for melanoma cell lines, with a zoom in on a cluster of compounds containing known cancer treatments. Each dot is a treatment represented as a 23-dimensional vector in which each dimension refers to the predicted $IC_{50}$ value in one of the 23 selected melanoma cell lines. Clinically approved monotherapies for *BRAF* V600E mutant melanoma are highlighted in bold. (**F**) $IC_{50}$ treatment space for a subset of brain and central nervous system cancer cell lines, where calcium channel blockers (zoom in) are clustered together and exhibit a differential effect. (**G**) To model cell line space with Prophet, each cell line is represented as a vector in which each dimension is Prophet's $IC_{50}$ prediction for one compound in that cell line. (**H**) Here we model the activity of EGFR inhibitor treatments in non-small cell lung cancer (NSCLC) cell lines. Each point is a cell line represented as a vector in which each dimension is the predicted $IC_{50}$ value for each treatment in that cell line for a set of NSCLC cell lines and EGFR inhibitors. The cell lines HC4006, NCIH1975, NCIH1650, HCC827 and CALU3 are known to be sensitive to EGFR inhibitors. Points are colored by predicted $IC_{50}$ sensitivity (Methods) to EGFR inhibitors. Lower $IC_{50}$ values imply higher sensitivity and higher cell death.

Prophet encounters challenges in outperforming the mean estimator in certain scenarios, such as predicting the effects of observed treatments in unobserved cell lines within SCORE, or forecasting the impacts of untested small molecules in known cell lines within LINCS. Further investigation revealed that the correlation between experimental replicates is a crucial metric for evaluating the predictability of their outcomes. We conducted a statistical analysis of the correlation between replicates (Table S4). This analysis revealed that, along the problematic axis, replicates of the same experimental conditions are less similar to one another than to the replicates of any other random condition. In such cases of noisy data with a strong batch effect the mean estimator is the best possible predictor.

## Prophet learns a universal landscape of biological experiments

Drug screens are limited in scope, either by the number of compounds tested or the variety of cell states. Hence, they provide a limited view of the profile of a compound, encompassing just the effects in the tested cell states. Including more cell states would provide a more robust and informative view of the pharmacological profile of a drug (*1*). Now, with Prophet, we can examine the treatment effect *in silico* across hundreds of cell lines at once. We represent each small molecule—or combinations of small molecules—as a vector in which each dimension is the Prophet-predicted viability value in each cancer cell line (Fig. 3a). This vector represents an "activity-based fingerprint" for each compound. In this high-dimensional representation, compounds with similar sensitivity profiles across cancer cell lines cluster together when viewed using a 2D representation (Fig. 3b). We can draw comparisons between these shared response

patterns for unbiased annotations of mechanisms of action (Fig. 3c), identification of new therapeutic candidates (Fig. 3d, e), or isolation of tissue-specific mechanisms (Fig. 3f, fig. S7a).

**Compounds which target pathways of cell proliferation.** We used Prophet to generate viability predictions for 8,620 small molecule treatments in 979 cell lines (Fig. 3a). We concatenated the 979 cell line values per drug to construct an activity-based fingerprint, which together form a viability-contextualized small molecule manifold (Fig. 3b). We focus on four distinct clusters in this space (Fig. 3c, fig. S6a) containing a total of 64 compounds that target specific regulators of cell proliferation: two from the MAPK signaling cascade (containing compounds targeting MEK and EGFR), and two from the PI3K/AKT/mTOR pathway. Two of the 64 compounds have never been seen in the training data (fig. S6b). With each of the 8,620 compounds tested in an average of 365 out of 979 cell lines, Prophet accurately completed 75.2% of the cell line $\times$ drug experimental space, grouping compounds in an unbiased manner after limited or no training examples.

Prophet's identified clusters are also biologically precise: after annotating the 64 compounds (Table S7, Supplementary Methods, fig. S6a), we found only one potential misclassification: the HSP90 inhibitor tanespimycin, which Prophet identifies as similar in effect to mTOR inhibitors(*42*). Additionally, we noted one misannotation: Ibrutinib, typically labeled as a BTK inhibitor, has been shown to bind to the EGFR mutant kinase in non-small cell lung cancer (NSCLC) (*43*, *44*). We suggest that Ibrutinib also has a similar effect in other cancer types. The joint clustering could be a pure effect of chemical similarity—however, we find compounds show considerable dissimilarity in chemical space (fig. S6c, Supplementary Methods). This indicates that the experiment representations learned by Prophet are driven by the phenotypic effects of small molecule treatments rather than their chemical structures, underscoring the utility of this manifold in capturing biologically relevant information which could not be identified using structural similarity alone.

**Therapeutic treatment candidates in *BRAF*-V600E mutant melanoma.** We wondered whether we can identify compounds which phenocopy existing, clinically approved treatments. To this end, we examined Prophet's predictions for four compounds already clinically approved for *BRAF* V600E-mutant melanoma. From the full 979 cell line viability space above, we create a subspace of 23 melanoma cell lines categorized as either *BRAF* V600E-mutant or -wild-type (Fig. 3e). 74% of these experiments were unseen in training. Such cancers respond well to various MEK and BRAF inhibitors (*45*–*47*). This subspace reveals a cluster of compounds containing three clinically approved monotherapies (*46*, *48*, *49*) and two combination therapies (*47*, *50*) which are highlighted in bold (Fig. 3e). 45% of the other compounds in this cluster have been indicated for *BRAF* V600E melanoma in *in vitro* experiments (Table S8), suggesting that the remaining compounds are relevant therapeutic candidates.
We further examine the predicted viability per cell line. Two out of 6 wild-type cell lines were on average sensitive to compounds in this cluster, compared to 9 out of 17 *BRAF* V600E-mutants (Fig. S8a). Among these, resistance to MEK or RAF inhibition despite the *BRAF* V600E mutation has been documented in certain cell lines, such as A2058's resistance to MEK inhibitor

PD0325901 (*51*), a behavior our model accurately predicts (fig. S8b) even though the current mechanism of resistance remains unclear.

**Predicting brain and CNS-specific monotherapies.** A similar *in silico* generated subspace of perturbation effects in brain and central nervous system (CNS)-derived cell lines reveals a cluster of neurological compounds, enriched for putative calcium channel blockers (p-value<$10^{-7}$, Fig. 3f, Supplementary Methods) which did not cluster when looking at global tissue responses (fig. S7c). Calcium signaling promotes growth and tumorigenesis of brain cancers (*52*). Recent studies demonstrate that calcium channel inhibition suppresses brain tumor growth, decreases the expression of oncogenes and increases expression of tumor suppressor genes, even proposing to repurpose existing drugs to treat brain cancers (*53*, *54*).

**Experimental space contextualized by RNA expression.** While the compound space detailed above was constructed using Prophet-predicted viability values, Prophet can also construct a compound space using other readouts, such as RNA expression learned from the LINCS dataset. We make predictions for BIRC5 expression after treatment in the same set of compounds and cell lines (fig. S7d), expanding the original by 5,846 small molecules and 922 cell lines. This new space captures patterns of BIRC5 expression after compound treatment (fig. S7e). This enables us to identify a set of MEK inhibitors affecting BIRC5 expression and reveals a new cluster of compounds acting on BIRC5 (fig. S7f). Although BIRC5 expression after perturbation has not been reported in the literature for many of these compounds, the cluster includes a survivin inhibitor and Auranofin, which has been reported to downregulate BIRC5 in osteosarcoma(*55*). Both of these compounds were not previously tested in LINCS.

**Identifying EGFR sensitivity in non-small cell lung cancer.** Applying drug activity fingerprint concept to cell lines, we can also generate a fingerprint for cell lines which we can use to investigate the response signatures of biological systems to a specific array of drugs. As an example, we identify non-small cell lung cancer (NSCLC) cell lines that react to EGFR inhibition. We subset the experimental space to NSCLC cell lines and compute an average EGFR inhibition sensitivity score per cell line (Fig. 3g, Supplementary Methods). *EGFR* mutant cell lines, with a constitutively active EGFR receptor, are often especially sensitive to EGFR inhibitors (*56*). Prophet identifies a gradient in the cell lines with respect to sensitivity to EGFR inhibitors, and is successful at scoring five known putative *EGFR* mutant or EGFR-sensitive cell lines as EGFR sensitive (H1975, H1650, HCC4006, HCC827 and CALU3) (*57*–*59*) (Fig. 3h).

## In vitro experimental design and validation

To further demonstrate the accuracy of Prophet's predictions, we experimentally validated a subset of the model's predictions. To do so, we used a setup with existing pre-built 384-well plates, each with 352 unique drug perturbations applied to 9 cancer cell lines (Table S9). In total, there were 16 pre-built plates. This setup represents a scenario in which limitations in experiment size constrain the compounds and cell lines that can be tested.

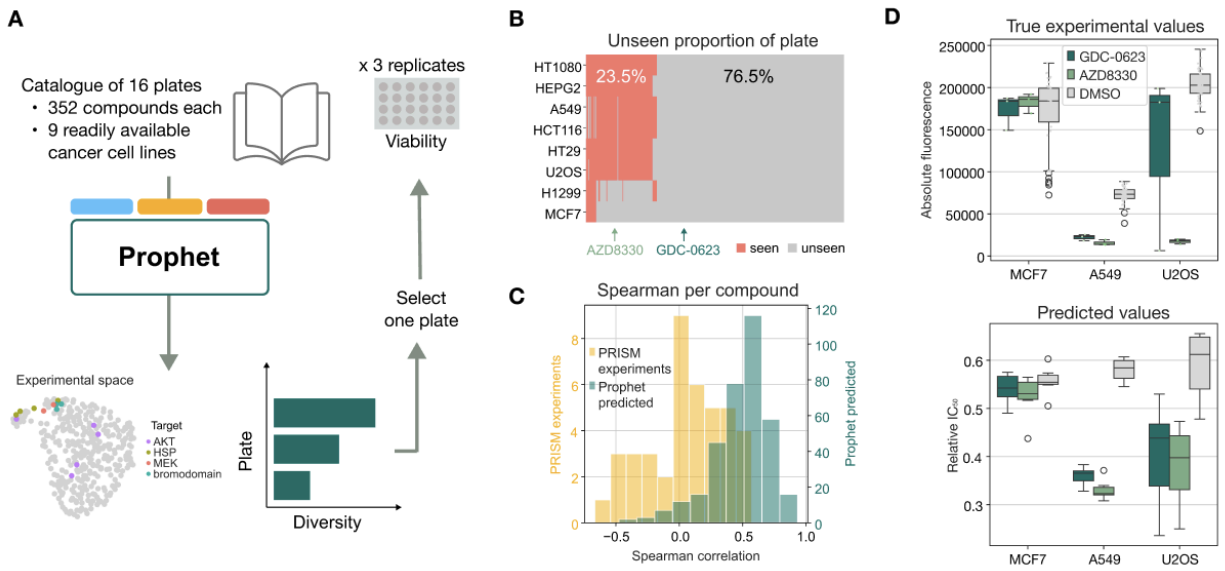


**Fig. 4. In vitro validation of selected Prophet predictions.** (**A**) Workflow accommodating the constraints of an in-house small molecule screening setup. (**B**) Proportion of the performed screen with completely new experiments. Cell line and drug experiments not seen by Prophet in any training dataset shown in gray. Experiments which were measured at least once, with any readout, are shown in red. Each column represents one compound. The two compounds also in panel D are highlighted. (**C**) Spearman correlation per compound of the real vs. predicted viability values across cell lines (teal, right axis) for the 352 compounds we experimentally validated. In yellow are the same correlations, but with the real viability values measured from PRISM. (**D**) Absolute CellTitreGlo fluorescence values from the in vitro experiments (three replicates) vs. Prophet's predicted values (6 ensembles). Boxes show the quartiles, and whiskers show 1.5 times the interquartile range from the box. Prophet provides tight uncertainty bounds for these predictions, and more loose uncertainty for the U2OS cell line.

Aiming to pick a plate with maximally diverse drug responses across cell lines, we used Prophet to make *in silico* predictions for the outcomes of the compounds in each plate in the nine available cell lines. We used the predicted effects on cell viability to score each plate in two categories—dissimilarity across compounds and similarity within clusters (fig. S9a, Supplementary Methods). We then tested the selected plate and compared the results to Prophet's predictions. Prophet successfully predicted the overall outcome for the entire plate, with a median of 0.53 Spearman correlation (Fig. 4c), despite not having seen any related data during training. Comparing the predicted viability scores for each compound of the selected plate with viability measurements in PRISM, the dataset with the best coverage of the compounds and cell lines out of which 8/9 cell lines and 41/352 compounds were also tested, resulted in a median of 0.07 Spearman correlation (Fig. 4c).

For the two known MEK inhibitors tested, GDC-0623 and AZD8330, Prophet correctly predicted no significant change in cell viability after treatment specifically in MCF7, but none of the other cell lines, with high certainty (Fig. 4d, fig. S9b). Our experimental validation confirmed this observation (Fig. 4d), showing that both MEK inhibitors produced no significant changes in viability in MCF7 cells, but significant changes in other cell lines. This finding is supported by

previous work documenting MCF7 resistance to some MEK inhibitors, including trametinib(*60*), but not other MEK inhibitors such as selumetinib and pimasertib(*61*). Prophet independently uncovers this highly specific resistance despite not previously seeing either compound in MCF7 cells in training, and not seeing GDC-0623 at all.

## Prophet predicts *in vivo* phenotypes

Prophet's capabilities extend to *in vivo* experimental settings. Saunders et al. 2023 (*28*) generated a single-cell transcriptomic atlas of the effect of 23 genetic knockouts—including combinations—on 1,812 developing zebrafish embryos across 5 timepoints (Fig. 5a). Their analysis includes using the single-cell transcriptomics data to compute relative cell type frequencies for each embryo (Fig. 5b) and comparing the genetic perturbations using similarities and differences in cell type proportion change. In this high-dimensional space, genetic perturbations clustered together based on the affected zebrafish developmental lineage. Given our framework, we asked if Prophet could learn these cell type proportion changes and make predictions about additional knockouts for specific lineages.

Given that this is the only *in vivo* dataset and that no common prior knowledge exists for it within our dataset compendium, we did not use any pretrained Prophet version, but trained a single Prophet model from scratch (Methods). We then used the predicted relative cell type frequencies to construct a zebrafish cell type proportion experimental space, extending it with predictions of single knockouts, and analyzed perturbation-dependent deviations relative to wild-type embryos (Fig. 5c).

Despite only 23 perturbations in the dataset, the cell type proportion experimental space learnt by Prophet qualitatively matches the one obtained by the authors in Fig. 2c of the original publication (Fig. 5d, top panel), where it can be seen that different gene knockouts provoke different responses in terms of cell type proportion in the zebrafish embryos. The gene knockout fingerprints learnt by Prophet group genes according to their involvement in the mesoderm, nervous system and neural crest lineages (Fig. 5d, bottom panel). We also run predictions for singleton knock outs of zebrafish transcription factors (TFs): for example, znf711 and znf740b from the zf-C2H2 family and dbpb, bach1b from the bZIP family, to produce additional candidates for further experiments.

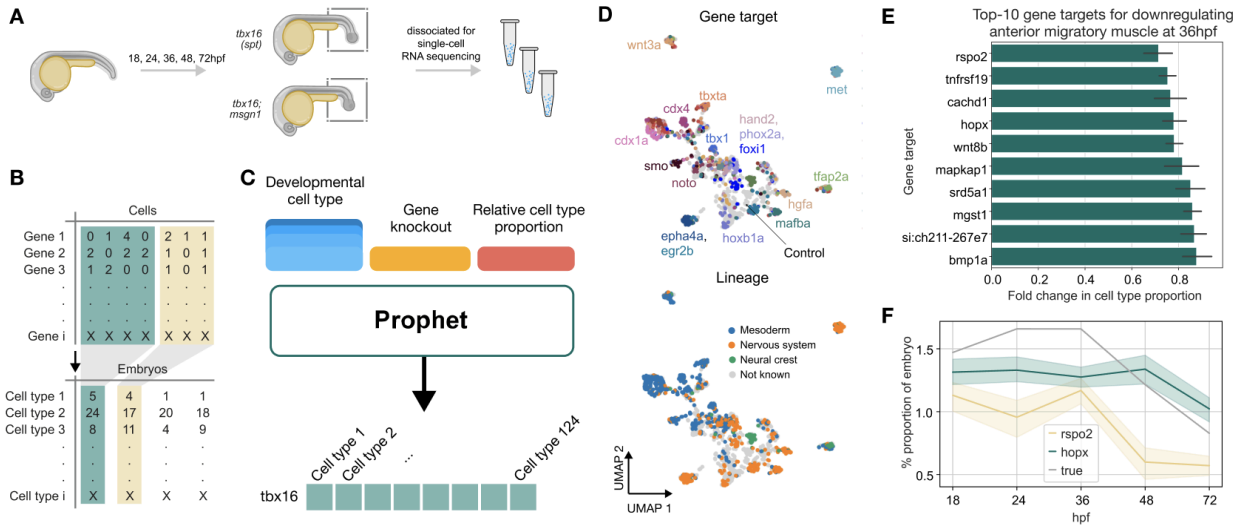


**Fig. 5. Prophet extrapolates relative cell type abundance in zebrafish development.** (**A**) Summary of the original experimental setup from Saunders et al. 2023. (**B**) How the single-cell x gene matrix was converted into cell type proportions per embryo. The latter is the phenotypic readout predicted by Prophet. Adapted from Saunders et al. 2023. (**C**) Schematic of how Prophet generates a complete experimental space *in silico*. Prophet predicts the cell type frequencies of one zebrafish embryo for each genetic KO combination across 124 cell types to construct a genetic KO "fingerprint". (**D**) Prophet's learnt experimental space of genetic KO "fingerprints".. Depicted are: combinatorial predictions which include a knockout seen in the original publication (top panel, colored by the seen knockout; bottom panel, colored by the lineage of the seen knockout) and singleton predictions for zebrafish transcription factors (gray) not seen in the original publication. Prophet's experimental space recapitulates the findings obtained by the authors in the original publication and the separation between genes involved in different developmental lineages. (**E**) Top-10 gene targets that suppress the appearance of the "anterior migratory muscle" cell type as predicted by Prophet, ranked by their average effect on the fold-change of the cell type proportion as compared to control. Rspo2 and Hopx are known regulators of mesoderm development, while Wnt8b and Bmp1a belong to families of genes also involved in the same lineage. (**F**) The predicted absolute cell type proportion of anterior migratory muscle at 5 timepoints for the top 1 and 3 predicted gene targets, with control as comparison.

As an example, we run Prophet to get predictions of gene targets that reduce the cell type proportion of anterior migratory muscle at 36 hours post-fertilization (hpf)(Supplementary Methods). We measure the strength of a knockout as fold change using the control embryo cell type proportion at 36hpf as a reference. Among the top 10 gene targets predicted (Fig. 5e), we find that Rsp02 and Hopx rank first and third, respectively; for both genes, previous studies describe their key role in mesoderm development (*62*); (*63*, *64*). Other genes in the top-5 (Bmp1 and Wnt8b) are part of families of genes also involved in mesoderm development(*65*–*67*). All ten of the Prophet-predicted strongest gene knockouts are not part of the original dataset. Prophet's predictions also report an uncertainty obtained from the variance of predictions of an ensemble of models (Methods). This informs experimentalists of Prophet's confidence in a prediction, allowing them to make an informed decision about whether or not to include high uncertainty TFs in a panel.

Finally, Prophet can also provide fine-grained predictions across timepoints by incorporating the time of the readout into the readout token. The proportion of anterior migratory muscle cells varies across time points given an F0 knockout in a zebrafish embryo in the genes Rspo2 and Hopx (Fig. 5f). These detailed temporal predictions can assist biologists to make more informed decisions about which subsequent experiments to conduct, such as focusing on particular stages for gene expression analysis or investigating downstream effects during critical developmental windows.

This example showcases how Prophet is flexible enough to extend beyond cell lines and *in vitro* experiments to *in vivo* phenotypes. In the future, Prophet could be provided with genetically orthologous data in human developmental models to bridge the gap between controlled laboratory experiments and complex *in vivo* systems, helping to translate findings from basic research into humans and potential therapeutic applications.

# Discussion

Prophet introduces a novel and scalable approach for phenotypic experimental biology, leveraging transfer learning to effectively predict outcomes across a diverse range of biological experiments. It significantly outperforms simpler architectures and the mean baseline in prediction on unseen experimental conditions (Fig. 2b, fig. S5c). Prophet learns a biologically meaningful experimental space that can be explored to generate new hypotheses. For instance, Prophet's precise predictions allow users to make screens 60x smaller while identifying the same number of hits (Fig. 2f). Activity-based compound fingerprints made possible by Prophet accurately recover known biological pathways, with predictions validating *in vitro*. Notably, Prophet's capabilities extend beyond cell lines to encompass predictions for *in vivo* settings (Fig. 5), laying the groundwork for systematic phenotypic drug discovery even in complex biological systems. These evaluations underscore Prophet's potential as a first step towards a foundation model for experimental biology, highlighted by its scaling properties, transfer learning capabilities, and performance on extensive datasets gathered.

Limitations of the current approach are, firstly, the simplified assumption of the three-axis experimental space (Fig. 1). There exist other elements that define an experiment, such as a temporal dimension or the dosage employed, as well as more subtle variations such as technical variation or protocol choice. Prophet considers the time point at which an experiment is measured to be a specification of the readout, without considering the ordinal nature of the variable. Drug dosage could be considered similarly, and these should add ordinal dimensions to experimental space as a topic of future development. Similarly, Prophet could take as input directly the genetic sequences or chemical structures of the treatments, without any preprocessing, or extend to antibody treatments with the appropriate representation and dataset. While Prophet can currently be applied to genetic perturbation datasets, fully leveraging its transfer learning capabilities will require incorporation of more extensive phenotypic datasets

of genetic perturbations. With future genetic intervention data and joint datasets, Prophet is posed to effectively translate knowledge between genetic screens and small molecule assays.

While Prophet successfully transfers learning between independent assays and phenotypes, a challenge that remains unsolved is how to transfer learning across scales—between *in vitro*, *in vivo* and patient assays. Mastering this transfer is crucial for improving clinical outcomes in drug discovery because more complex functional assays, using organoids, patient samples, or *in vivo* experiments can often be limited in size. If supplied with this data, the Prophet architecture and framework would enable more accurate predictions by transferring information from scalable *in vitro* perturbation screens, up the chain of complexity to different biological contexts.

Prophet represents a new concept in phenotypic drug discovery, demonstrating that transfer learning between independent assays is not only possible but also enhances results. Its analysis capabilities and general architecture position it as a step towards establishing a foundation model for perturbation biology: a comprehensive unified model that integrates diverse types of information under a single learning paradigm. Prophet's methodology allows it to be used as a backbone for a lab-in-the-loop system to propose the next set of perturbational experiments, guiding phenotypic drug discovery and navigating the vast experimental space. These attributes make Prophet a powerful tool for driving progress in biological research and experimental design.

# Materials and Methods

### *External datasets*

We selected datasets with at least 100 clearly labeled cell states or treatments. In most cases, we used as the predictand the least processed values post-batch correction generated by the original authors. The contents of each dataset are detailed in Table S1.

All datasets then underwent the following: where the original authors did not collapse replicates, replicates were averaged. Experimental outcome values were scaled from 0 to 1. Cell lines, treatments, and readouts which do not exist in the input lookup files were excluded from both training and test. As small molecules were represented using their SMILES strings, small molecules without SMILES strings or with ambiguous SMILES strings were filtered out.

### *Baseline models*

The baseline models we employ are linear regression, a multilayer perceptron (MLP) and Random Forest. As some experiments are combinations of treatments and the baselines are not permutation invariant, all datasets are duplicated and the treatments flipped. This means that an observation $X_i = \{X_{env_i}, \{pert_{1_i}, pert_{2_i}\}, X_{phe_i}\}$ is also seen as $X_i = \{X_{env_i}, \{pert_{2_i}, pert_{1_i}\}, X_{phe_i}\}$. For the linear regression setting, L1 and L2 regularization are tested but just the best results are reported. Random Forest runs with 100 estimators, minimum samples per split=2 and without capping the max length.

### *Benchmarks of performance evaluation*

We distinguish between two distinct evaluation scenarios: unseen treatments and unseen cell states. While both are similar, they differ in their respective axes of the experimental space. For unseen treatments, we construct a hold-out test dataset comprising entirely new treatments. For combinatorial datasets including zebrafish, this means any combination including the unseen treatment is placed in test. For unseen cell states, we construct a hold-out test dataset comprising experiments in entirely new cell states. In both scenarios, the test dataset represents 20% of the total treatments or cell lines, while the validation dataset encompasses 10% of them.

Importantly, in the case of fine tuning a pretrained Prophet model, we ensure that there is no data leakage or any data split issue confounding the performance by maintaining consistent train and test splits between the large pre-trained model and the assay-specific models.

Additionally, we examine the model's performance as it is exposed to progressively more treatments or cell lines. Specifically, we assess Prophet's performance with 80%, 50%, 30%, 20%, 10%, and 5% of the treatments or cell lines included in the training dataset.

### *Overview of Prophet*

Prophet considers a perturbational dataset of $N$ experiments $D = \{(X_i, Y_i)\}_{i=1}^{N}$ where $X_i$ represents the experiment carried out and $Y_i \in R^K$ is the $k$-dimensional functional outcome resultant of the experiment. Each experiment $X_i$ is modeled as a set $X_i = \{X_{env_i}, X_{pert_i}, X_{phe_i}\}$ conformed by the environment where the experiment is carried on ($X_{env_i}$), the set perturbations applied ($X_{pert_i} = \{pert_{1_i}, pert_{2_i}, ..., pert_{n_i}\}$) and the phenotype that is measured ($X_{phe_i}$). The goal of Prophet is to learn a function that maps a novel experiment $X_j$ to a functional outcome $Y_j$.

Given a experiment description $X_i = \{X_{env_i}, X_{pert_i}, X_{phe_i}\}$, Prophet first projects all representations into a common token space $Z_i \in R^d$, where each token represents an element of the experiment, using the tokenizer networks. This learning process involves non-linear transforming, using feed-forward networks, prior representations if available; otherwise, it resorts to transforming simple one-hot encodings. Hence, $Z^0_i = \{Z_{env_i}, Z_{pert_i}, Z_{phe_i}\}$, where $Z_{env}$, $Z_{pert}$ and $Z_{phe}$ are the outcome of the tokenizer networks. In the case of the set of treatments applied $Z_{pert}$, <PAD> tokens are appended if needed to ensure a consistent length in all observations. Subsequently, we prepend a learnable class token <CLS> into the sequence of experiment-related tokens, resulting in the augmented set $Z^0_i = \{Z^0_{CLS}, Z^0_{env_i}, Z^0_{pert_i}, Z^0_{phe_i}\}$. The set $Z^0_i$ is then fed into the transformer, where the combination of different feed-forward networks, residual connections and self-attention modules produce an output representation $Z^L_i$ at the final layer $L$ of the transformer.

The resulting final state of the learnable <CLS> token $Z^L_{CLS}$ serves as a comprehensive representation of the experiment. This representation is employed for predicting the functional outcome $Y_i$ through a function $f: R^d \rightarrow R^K$. The entire network is trained end-to-end to minimize the MSE error between the predicted experimental outcome $\widehat{Y}_i$ and the real experimental outcome $Y_i$.

### *Tokenization*

We call tokenization the process of projecting elements of the set $X_i = \{X_{env_i}, X_{pert_i}, X_{phe_i}\}$, that represents a experiment, into a common token space to get representations $Z_i \in R^d$.

For each of these elements, a specific tokenizing function $f: R^p \rightarrow R^d$ is learnt where $p$ is the dimensionality of the initial representation and $d$ is the dimensionality of the token space. These functions project each element of the experimental set into the common token space that serves as input for the Transformer. In the case of the treatments, they undergo a different transformation $f$ depending on whether it is a small molecule or genetic perturbation. All the transformations are modeled as feed-forward networks.

$$Z_{env_i} = f_{env}(X_{env_i})$$
$$Z^i_{phe} = f_{phe}(X^i_{phe})$$
$$Z^i_{pert} = f_{CRISPR}(X^i_{pert})$$
$$Z^i_{pert} = f_{DRUG}(X^i_{pert})$$

***Cell state representations***

The most important determining factor in choosing cell state representations is the size and therefore ability to allow for unseen cell state predictions. For example, choosing a representation of cancer cell lines which only spans 100 cell lines would limit predictions to those 100 cell lines. As such, we used bulk RNA-seq from the Cancer Cell Line Encyclopedia (CCLE) (*31*) from Dependency Map, which covers 1406 cell lines. Further details can be found in Table S1.

For the one perturbation dataset performed in zebrafish embryos, we consider each cell type as a different tissue. To do so, we used the pseudo-bulked expression of the cell type in a control embryo across all time points as the cell state representation.

To reduce training time, we reduced the dimensionality of both sources using PCA to 300 components.

***Treatment representations***

All assays examined in this study involve perturbations induced by either CRISPR or small molecules. Hence, the prior representations of each class of perturbations are different to gather the relevant aspect of each. In both cases, however, the different representations are concatenated.

For CRISPR treatments (including KO, interference, and activation), we use the target gene to represent the treatment. We further go on to use two orthogonal sources of information to represent a target gene - its co-expression patterns across cell lines, and its coding sequence. Co-expression pattern representation is derived from CCLE. The coding sequence representation is derived from Enformer embeddings for the whole transcription sequence.

Enformer embeddings are derived from the second to last layer of the model. The genomic coordinates of GRCh38 (hg38) were extracted using the BioMart Data Mining tool. In this process, we selected the Human genes GRCh38.p14 dataset and only included genes with an HGNC symbol ID. For genes with multiple transcripts, we retained the coordinate with the smallest Transcription Start Site (TSS). Subsequently, we retrieved the corresponding gene sequences from the GRCh38 human reference sequence produced by the Genome Reference Consortium.

For small molecules, we use the canonical SMILES string to represent the treatment. We further go on to use different representations of molecular structure - RDKit2D(*32*) molecular graph embeddings and Morgan fingerprint embeddings(*33*). RDKit embeddings produced a 200-dimensional vector, while the Morgan fingerprints were calculated as 1024-dimensional vectors with radius fixed at 3. Specifically, we use 195 dimensions that represent the RDKit2D fingerprints and 1024 dimensions for the Morgan fingerprints. In cases in which two datasets used the same common name for a compound but attached different SMILES, we appended a suffix to the compound name (e.g. imatinib_GDSC and imatinib_PRISM) to differentiate between the two.

### *Phenotype representations*

Phenotypes are typically learned from scratch without any prior knowledge-based representation, primarily due to the challenge of identifying a meaningful prior that accurately reflects the phenotype. Hence, phenotype representations are modeled as learnable embeddings and developed during training.

### *In vivo* phenotype prediction

In the case of the cell type proportion in the zebrafish dataset, we do not use the pretrained Prophet version since it is a completely different organism than the ones in the pretraining corpus data and the data representations fundamentally change.

In this case, we use a timepoint representation which is the vector of cell type proportions of the control embryos at that timepoint. The cell state embedding is constructed by subsetting to control embryos, computing the top 300 highly variable genes, and then taking the pseudobulk RNA expression of the cell type. The treatment embedding is constructed from the 300 component PCA of the transposed cell x gene matrix of the control embryos, exactly as was done for the human CRISPR treatment embedding. Finally, each readout is the cell type proportion after a given F0 knockout at a given time *t*. To ensure no data leakage in evaluation, the cell type proportions of the control embryos are never used in test or validation.

### *Model training*

Prophet is efficiently trained in a single Nvidia A100 40GB GPU. We train the model using float32 precision. We use the AdamW optimizer(*68*) with $\beta_1 = 0.9$ and $\beta_2 = 0.999$, weight decay of 0.01, and dropout of 0.1. The batch size is 4096 in the case of the model training across all assays, while it depends on the size of the dataset in the case of the smaller experiments. The learning rate increases linearly until 1e-4 with a linear warmup of 10,000 steps. After the warmup, a cosine decay regime(*69*) is applied. All weights are initialized using Xavier initialization(*70*) with default parameters, while the bias terms are initialized to 0. It is run with an early stopping with patience of 20 epochs. During pre-training, it tracks the average R2 performance in the validation set per assay. However, we have noted that model selection after a sufficient number of pre-training epochs does not change results significantly. For fine-tuning, the early stopping tracks the dataset-specific R2 in the validation set. During fine-tuning, all parameters of Prophet are updated during 20.000 steps with a batch size of 256 observations.

The pre-trained Prophet model consists of 8 transformer encoder units with 8 attention heads per layer and a feed-forward network size of 1024 to generate a 512-dimensional embedding of the experiment. However, the Prophet models trained from scratch in specific datasets are much smaller (2-4 layers and 1-4 attention heads). This is because using a large model in dataset-specific cases resulted in significant overfitting (Supplementary Methods).

***Uncertainty***

With the exception of zebrafish development, we compute an uncertainty value for Prophet's predictions by training six individual models on all datasets. Each model maintains the same test set, but the validation set (and therefore the training set) is selected using an arbitrary seed. As Prophet generally exhibits very stable behavior when learning representations robustly represented in the data, the variance of these predictions can be used as a measure of the uncertainty of the model.

We take a similar approach to computing uncertainty in zebrafish development, but because the zebrafish dataset has no overlap with any existing dataset, we compute uncertainty using different test and validation genes with a model trained on only this dataset.

# Code and data availability

The code for Prophet can be found at https://github.com/theislab/prophet. Data obtained from in vitro validation can be found at https://figshare.com/articles/dataset/Prophet_viability_plate_xlsx/26516002.

# Acknowledgements

Madeleine Duran and Cole Trapnell for advice on zebrafish analysis and the list of annotated zebrafish transcription factors. Paul Bertin for suggesting the use of an accuracy score for biological interpretability, additional potential model architectures, and innumerable other helpful insights. Szen Toh for making suggestions for both datasets and embeddings, helping coordinate with the Garnet lab which authored so many of these datasets, and commenting on the initial draft of the manuscript. Tim Treis for invaluable advice on the cell painting datasets.

Luke Zappia for helpful and extensive commentary on the figure designs. Stefanie Brandner for excellent technical assistance in the small molecule library experiment. Riccardo Trozzo, Ekaterina Zhigalova, and Juan José Montero Valderrama for biological advice on CRISPR perturbations. Stefan Peidli, Tessa Green, and Camille Fourneaux for advising on content and clarity in the initial draft of the manuscript.

This work was co-funded by the European Union (ERC, DeepCell - 101054957) and supported by the Chan Zuckerberg Initiative Foundation (CZIF; grant CZIF2022-007488 (Human Cell Atlas Data Ecosystem)). F.J.T. consults for Immunai Inc., Singularity Bio B.V., CytoReason Ltd, Cellarity, Curie Bio Operations, LLC and has an ownership interest in Dermagnostix GmbH and Cellarity.

## Author contributions

A.T.L. and Y.J. conceived the study with the help of F.J.T., and designed all experiments; Y.J. and A.T.L. contributed equally and have the right to list their name first in their curriculum vitae; S.K., Y.J., and A.T.L. curated the datasets; A.T.L, Y.J, and X.Z. wrote the model; Y.J., A.T.L, X.Z., and Z.Z. prepared and wrote the initial draft of the manuscript. I.R., J.T., and K.H. performed the in vitro experiment. F.J.T. supervised the project. All authors read, corrected, and approved the manuscript.

# Supplementary

## Supplementary Methods

**Prophet-individual training**

Each Prophet-individual model has been optimized after an extensive hyperparameter search. They have from 2 to 4 transformer blocks and from 1 to 4 attention heads, depending on the specific dataset (Supplementary Table 7). The final models are selected according to an early stopping selection criteria that tracks the R2 score in the validation set, with patience of 20 epochs.

**Alternative cell line and treatment embeddings**

For the treatment embeddings, different prior-knowledge based options were tested, such as generating embeddings from STRING(*71*), Perturb-Seq datasets (*72*, *73*), gene2vec(*74*), GO pathways (*75*, *76*) or transposing the CELLxGENE(*77*) matrix and taking PCA to obtain a low dimensional gene representation(*78*).

Regarding the cell lines, the copy number variant information provided by the Cell Model Passports(*77*, *78*) was also tested.

None of those options provided better results than the presented settings.

**Computing a hit and AUPRC using in silico predictions**

We define a hit as predicted by Prophet to be a compound for which the fold-change in the readout of interest increases or decreases by 2x as compared to DMSO in the same cell line. For example, if A549 has a relative $IC_{50}$ of .9 to DMSO, only compounds predicted to have a relative $IC_{50}$ of .45 or lower would be a hit. We compute the area under the precision-recall curve using sklearn's `PrecisionRecallDisplay` with the true binary indicator as the 2x threshold computed on real values (or 4x where indicated in the text), and the *in silico* predicted fold-change values as the prediction vector. These statistics are translated into a 60x reduction in the number of necessary experiments by assuming an experiment in which the user assays the small percentage of compounds predicted to be positives by Prophet, and compares the number of hits obtained to an assay of the same size but in which random compounds from GDSC were chosen.

**Constructing activity fingerprints**

To construct a small molecule viability embedding space, we use the predicted $IC_{50}$ values of 979 cell lines as the vector representation for each compound. Cell lines were selected from those seen at least once in any dataset, and compounds were excluded if we *only* had an InChi key hash or Broad ID, such that we could not biologically interpret their use or activity, nor could we order them for further experiments. We then co-opted the default scanpy(*79*) workflow to normalize predicted values per cell line, compute PCA, neighbors (Euclidean distance with 15 neighbors), and finally UMAP(*80*).

For the embedding space constructed from chemical structure alone, the RDKit and Morgan fingerprint embeddings used to train the model were used as the vector representation for each compound. We subset the 156k available embeddings to the same 8620 compounds which are also displayed in Fig. 3b and preprocessed them in the same way.

While one could examine the experimental space learned by single-assay Prophet models, exploring the space of the Prophet model trained across a diverse array of experiments offers additional advantages. The larger model benefits from being trained on a more comprehensive dataset (Fig. 2), which enables it to capture a broader range of experimental conditions, leading to more robust and reliable predictions. Hence, its experimental space is better positioned to provide nuanced and contextually accurate insights, making it a more powerful tool for guiding biological research and experimental design. On top of that, it is a single model able to span different phenotype-specific experimental spaces, instead of multiple models creating multiple spaces.

**Metadata annotation of small molecules and cell lines**

Compounds were annotated with putative mechanisms of action and protein targets from DepMap (*81*, *82*) using their common name. Common names were retrieved from the respective metadata for each dataset. The following compounds which could not be automatically annotated and for which the name of the compound did not make the mechanism of action abundantly obvious were manually annotated in Supplementary Table 7.

**Enriched mechanism of action annotation**

DBscan(*83*) is an unsupervised community detection algorithm that identifies clusters according to the density, and importantly does not assign all points to a community. We use the pertpy (Heumos et al. 2024) implementation of DBscan to identify clusters using the following parameters: min_sample = 5, eps = 0.25. The outlier cluster and the largest cluster, representing non-effective drugs, were excluded from the analysis. Clusters in which at least 33% of compounds are annotated with the same mechanism of action or protein target are considered enriched for that pathway. P-values for this enrichment were determined by Fisher exact test.

**Computing cell line sensitivity to a set of compounds**

A cell line's sensitivity to a single compound is defined as

$$Sensitivity_{Cell\ Line,\ Compound} = \frac{IC50 - \overline{IC50}}{SD_{IC50}}$$

Where $\overline{IC50}$ and $SD_{IC50}$ are the mean and standard deviation of the $IC_{50}$ values across all 8620 compounds. We assume that most compounds will not have an effect on the cell line.

A cell line's sensitivity to a class of compounds is therefore:

$$Sensitivity_{Cell\ Line,\ Compounds} = \sum_{i=0}^{i=n} Sensitivity_{Cell\ Line,\ i}$$

**Scoring sets of 352 compounds for diversity**

As nine cell lines were prepared to be treated in triplicate in 384-well plates at 5 μM concentration and incubated for 72hrs (Methods, Supplementary Table 10), we used Prophet to predict the *in silico* outcome of each of these plates and scored them for diversity. First, we scored each plate for how negatively correlated compounds were amongst themselves. Second, we computed communities of compounds using leiden and then took the average of correlations within a community. From these, we prioritized a plate with low correlation across compounds but high correlation within communities in order to optimize for compounds which produced different effects across cell lines but also produced similar responses to other compounds on the plate, resulting in plate 16 (fig. S9a).

**Cell Lines and Culture Conditions**

Cell lines were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Thermo Fisher Scientific, 41966-029) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher Scientific), 1% Penicillin-Streptomycin (Thermo Fisher Scientific), and 1% non-essential amino acids (MEM NEAA, Thermo Fisher Scientific). The cells were maintained at 37°C in a humidified atmosphere with 5% CO2.

The following cancer cell lines were used for the screening: human lung carcinoma cell line A549, human hepatocyte carcinoma cell line HepG2, human colon carcinoma cell line HCT116, human fibrosarcoma cell line HT-1080, human colon adenocarcinoma cell line HT29, human T-cell leukemia cell line Jurkat, human breast cancer cell line MCF7, and human osteosarcoma cell line U2OS.

The following cell lines were provided as gifts: HCT116 from Götz Hartleben, HT29 from Joel Schick, Jurkat from Daniel Krappmann, and MCF7 from Martin Selmansberger. A549, HCT116, HepG2, HT-1080, and U2OS cells were purchased from ATCC.

**Compound Screening**

For the screening, 500 cells per well were seeded in 80 µl of medium in 384-well microplates (CulturPlate, PerkinElmer). After 24 hours, the cells were treated with compounds (0.4 µl per well, 5 µM final concentration) or DMSO alone as a control. The HTS platform system, consisting of a Sciclone G3 Liquid Handler (PerkinElmer), was used for plate and liquid handling. The screening plate used in this study is part of a small-molecule library with known modes of action (dissolved in DMSO, 1 mM stock solution). After 72 hours of incubation, cell viability was measured using 20 µl of CellTiter-Glo 2.0 Reagent (Promega). Luminescence signals from the CellTiter-Glo assay were detected using the EnVision 2104 Multilabel Plate Reader (PerkinElmer). The luminescence signals from the DMSO control wells were used as the baseline, representing 100% activity.

**Constructing the genetic KO experimental space for zebrafish development**

Because the original experiment contained combinatorial perturbations, it was possible for us to also make combinatorial predictions during inference. However, the computational cost of the pairwise combination of all genes would have been immense, and therefore we made predictions for 1) the original 23 genetic KOs from Saunders et al, in all pairwise combination with 200 randomly selected transcription factors (Fig. 5c) and 2) the single genetic KOs of 2.1k transcription factors, as provided to us by Madeleine Duran from the Trapnell lab (Fig. 5d).

## Supplementary Figures

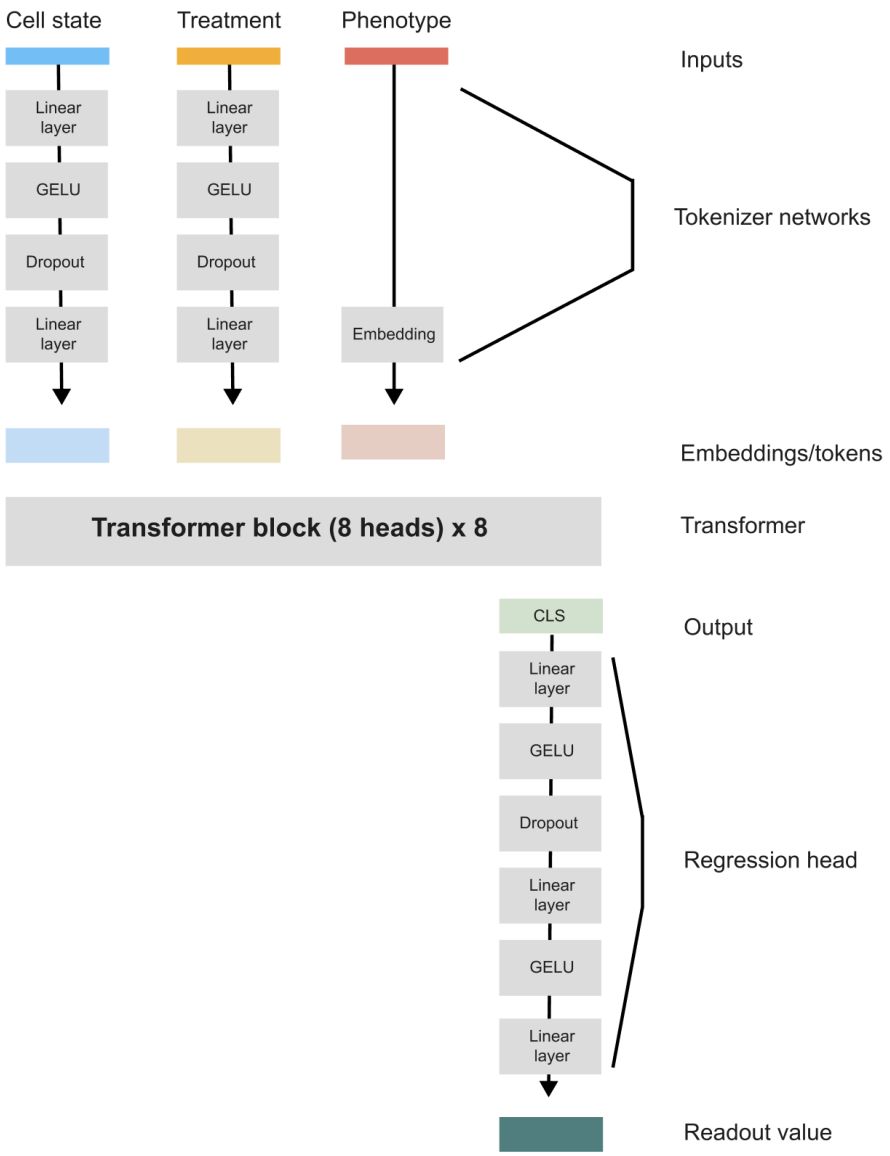


**Fig. S1. Prophet's architecture.** Prophet is a transformer-based model consistent of axis-specific MLPs that act as tokenizer networks, 8 transformer blocks with 8 attention heads each, and a MLP as a regression head that outputs the readout value.

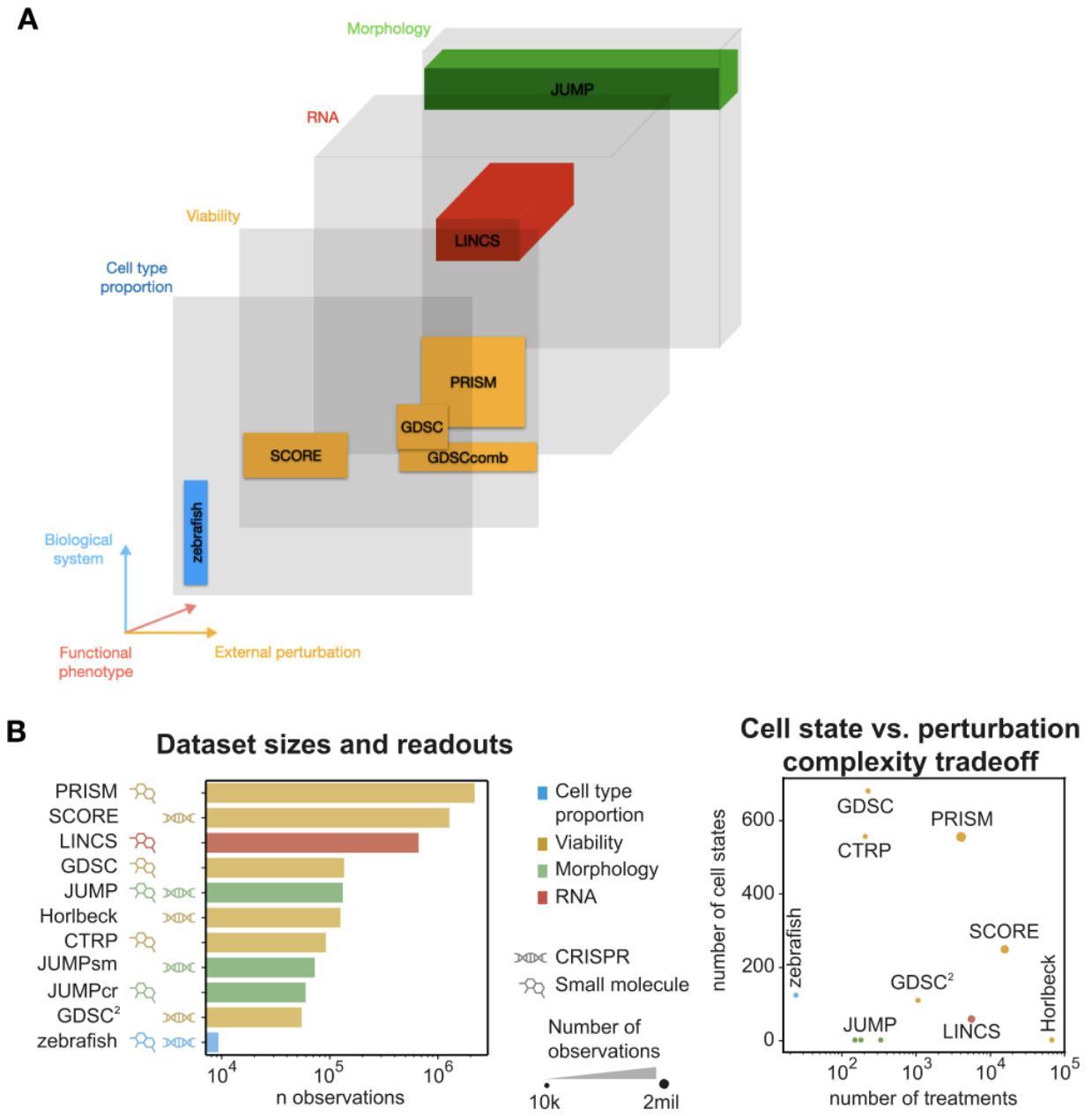


**Fig. S2. Overview of the datasets.** A) Approximate representation of the coverage of the dataset compendium in experimental space. We gathered 9 perturbational datasets that encompass up to 4 different phenotypes: $IC_{50}$/viability, RNA measurements, cell type proportion and cell morphology, with varying amounts of overlap. In total, there are more than 4.7 million experiments. B) The datasets vary both in number of cell states and number of interventions, and include drug and genetic perturbation. Not indicated: $GDSC^2$, Horlbeck, and zebrafish are combinatorial.

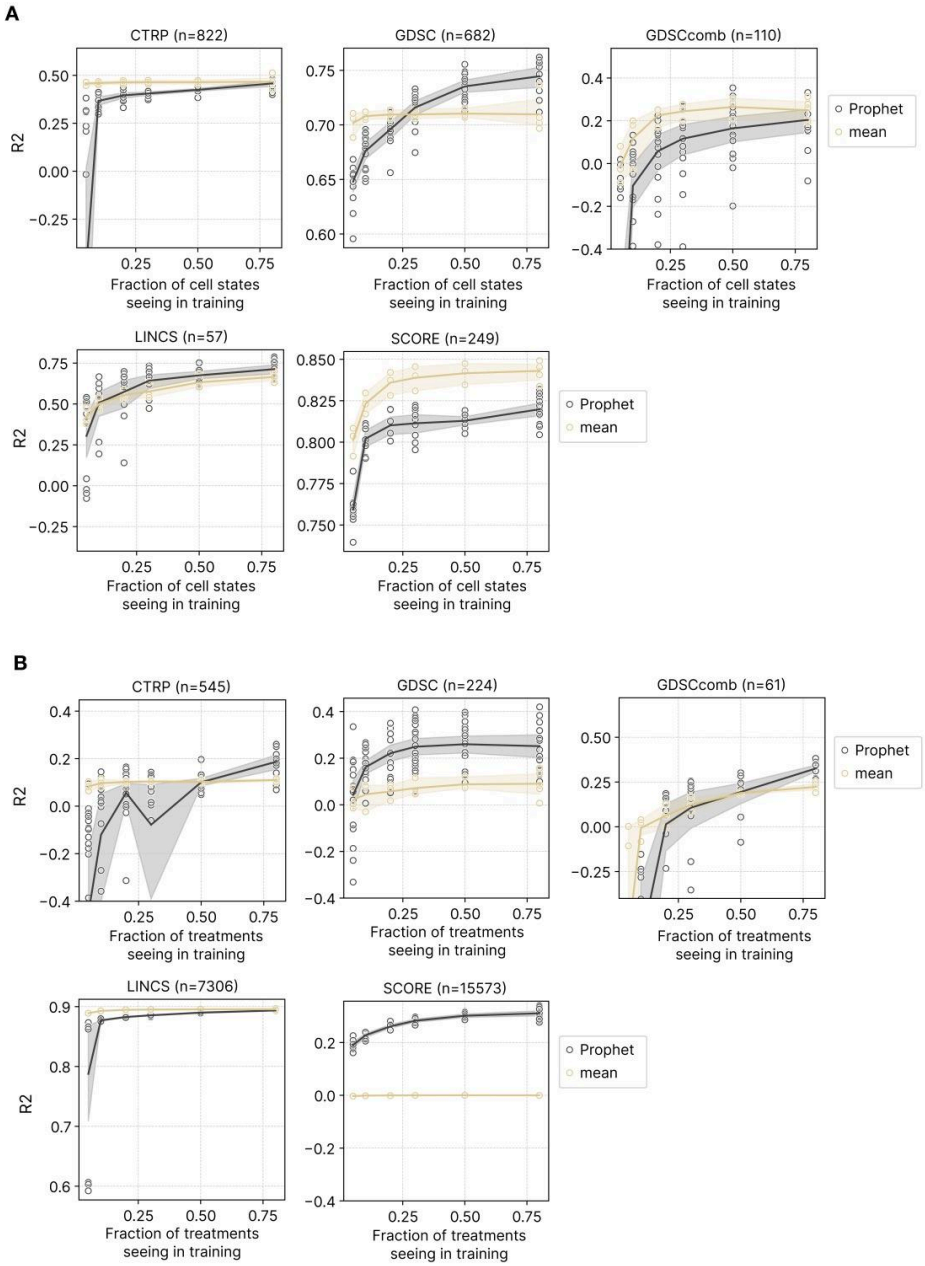


**Fig. S3. Prophet-individual models' performance as a function of the number of conditions seen during training.** The more the experimental space is mapped during training, the better Prophet's performance. However, the concrete pattern is dataset dependent as different datasets capture different amounts of variation. A) Performance of Prophet-individual models along the cell line axis. B) Performance of Prophet-individual models along the treatment axis.

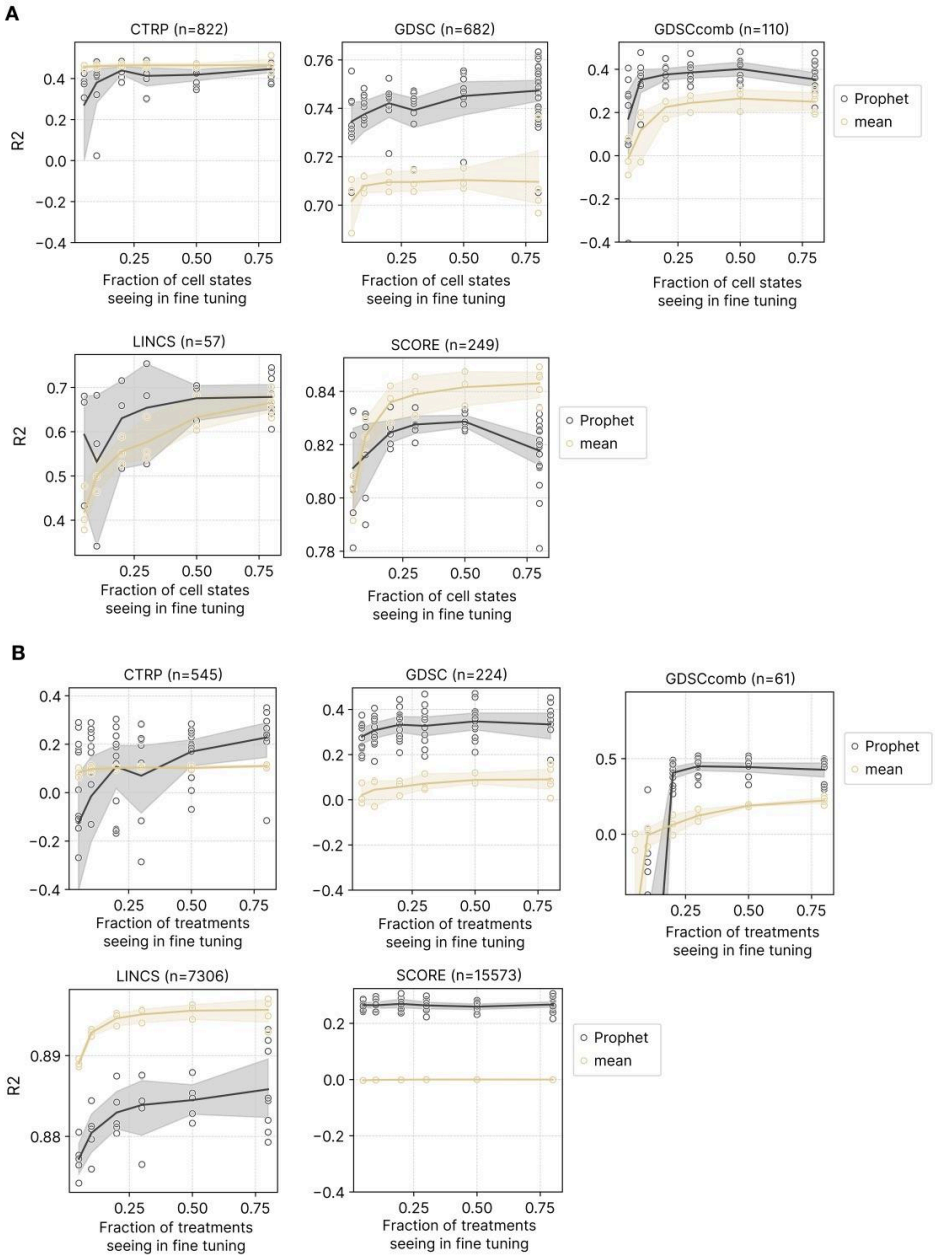


**Fig. S4. Prophet models' performance as a function of the number of conditions seen during fine tuning.** It can be seen that performance is more robust when fine tuning. As a remark, all models here are fine tuned for just 20,000 steps, independently of the size of the dataset. A) Performance of Prophet along the cell line axis. B) Performance of Prophet along the treatment axis.

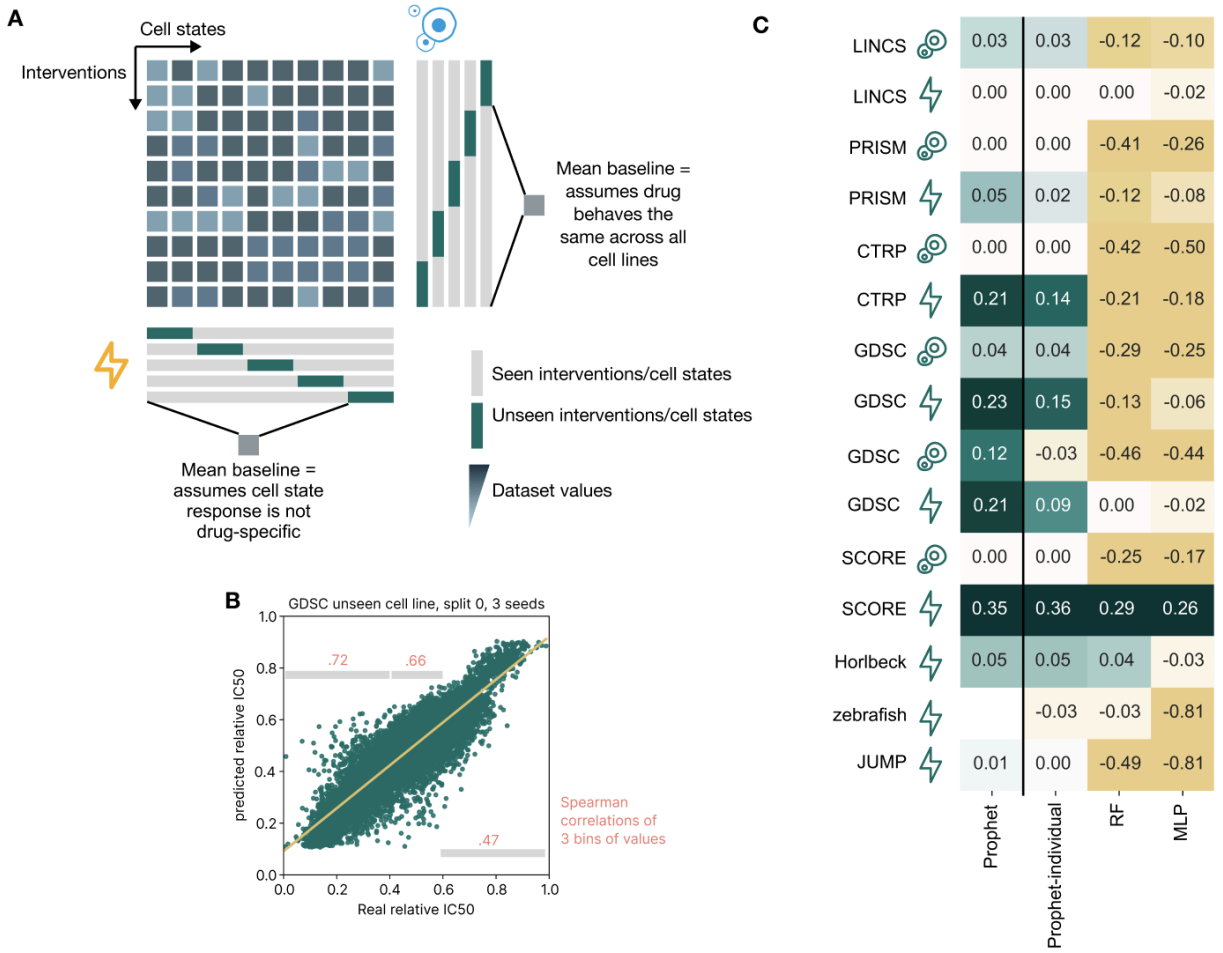


**Fig. S5. Mean baseline and evaluation schema.** A) Schematic of how 5-fold cross-validation is constructed for a single seed, resulting in ten runs. This is repeated for 3 seeds. The grey and blue tiles each represent one value in the experimental matrix of a dataset. The mean baseline estimates that the effect of an intervention in an unseen cell state is the average value of that intervention in all previously seen states. Conversely, it estimates the effect of an unseen intervention on a cell state is the average of all seen interventions in that same cell state. Naturally, it changes depending on the test set used. B) Spearman correlation of three different bins of the data, for one split in GDSC unseen cell line prediction across three seeds. Values are binned according to the real $IC_{50}$ values, with the cutoffs being the 40% and 60% quantiles. C) The same performance table as in Fig. 2b, but reporting Spearman correlation instead of R2.

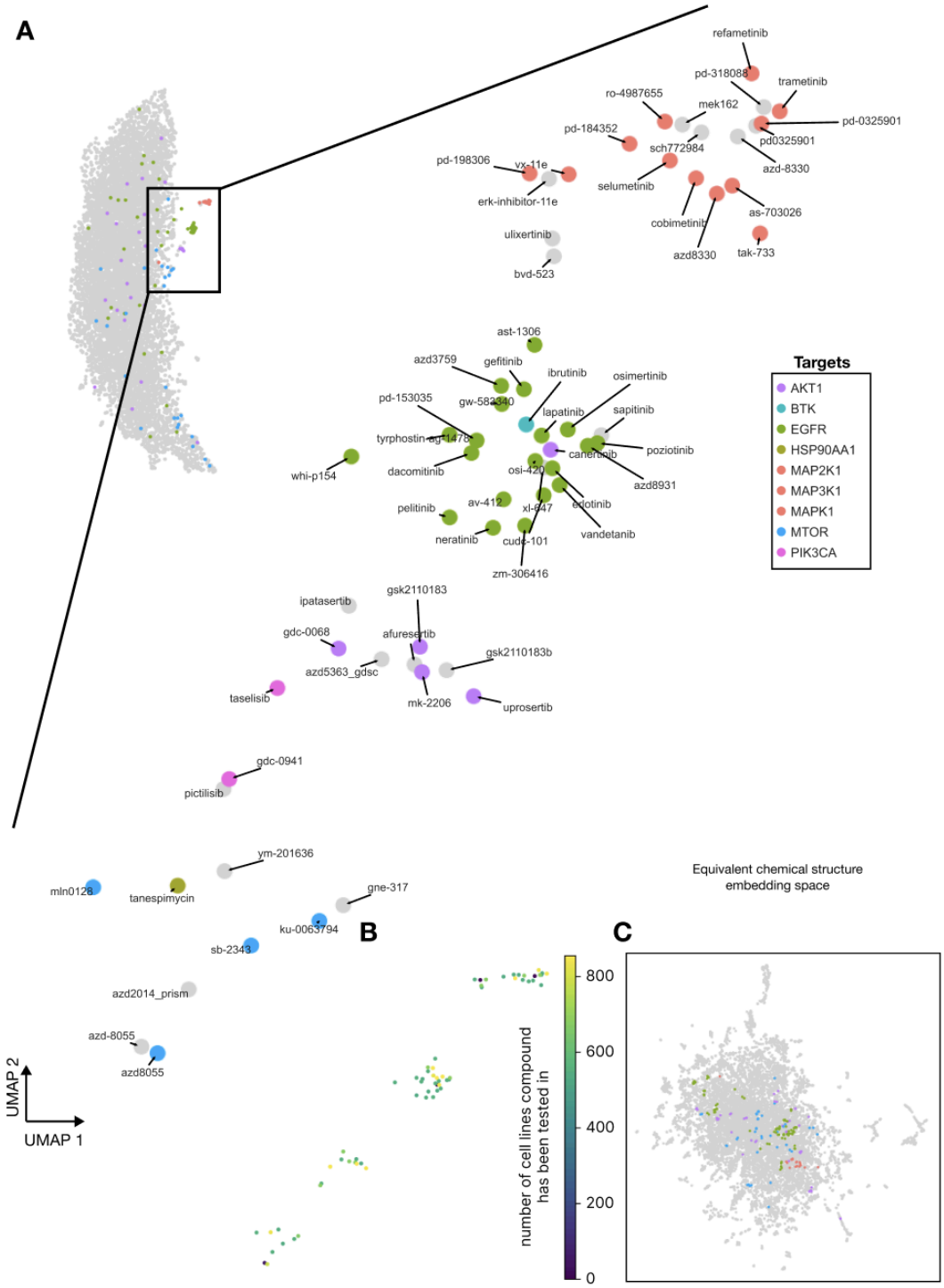


**Fig. S6. Detail of the clusters found in the viability Prophet experimental space.** A) Drugs are annotated by their automatically fetched annotations from pertpy. Compounds in grey did not have a putative target annotation in the database, but may still be annotated by existing literature. B) The number of unique cell lines in which a compound has been tested, for the 4 zoomed-in clusters. The fewest is 0 and the most is 829. C) Chemical structure embedding for the same set of compounds as shown in A, with the same 4 clusters colored.

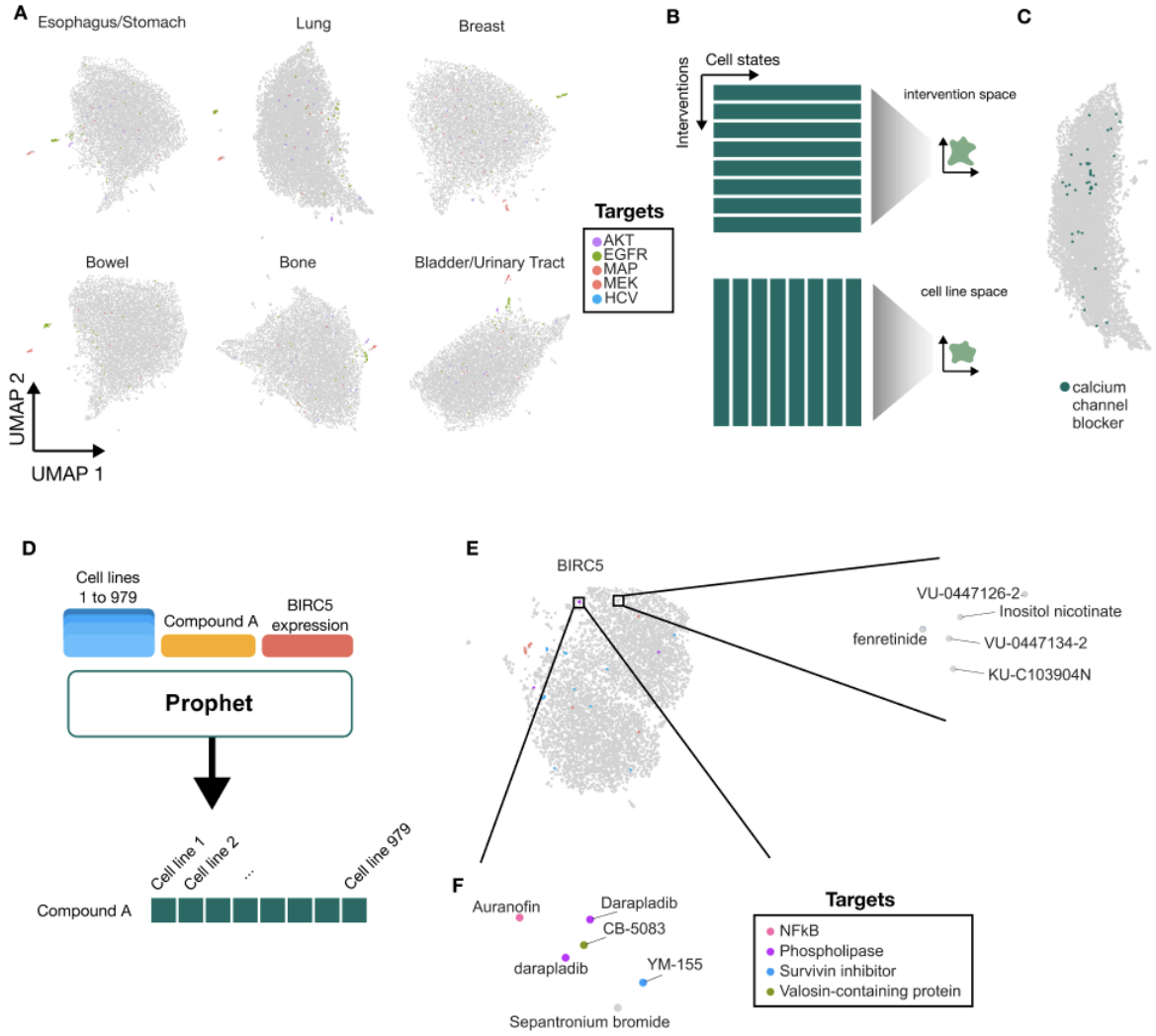


**Fig. S7. Experimental space construction and additional examples.** A) The 8620 compounds depicted in Fig. 3b, but subset to cancer cell lines of different tissue origins. Different tissues demonstrate different response patterns to compounds. B) Graphical depiction of how cell line spaces and intervention spaces are generated. Each teal rectangle represents a vector which is embedded as a 2D point using UMAP. C) Calcium channel blockers in the experimental space shown in Fig. 3b. Calcium channel blockers do not have a cohesive activity fingerprint across all tissue types, in contrast to when subset to brain and CNS-derived cancers. D) Schematic of how the BIRC5 experimental space is constructed. E) The experimental space of BIRC5 expressed across 979 cell lines after treatment with 8620 compounds. This is the BIRC5 RNA expression space equivalent of Fig. 3b, which depicts $IC_{50}$ space instead. A potentially interesting cluster of unknown compounds is annotated (right). C) A cluster of compounds with some putative regulators of BIRC5 (survivin).

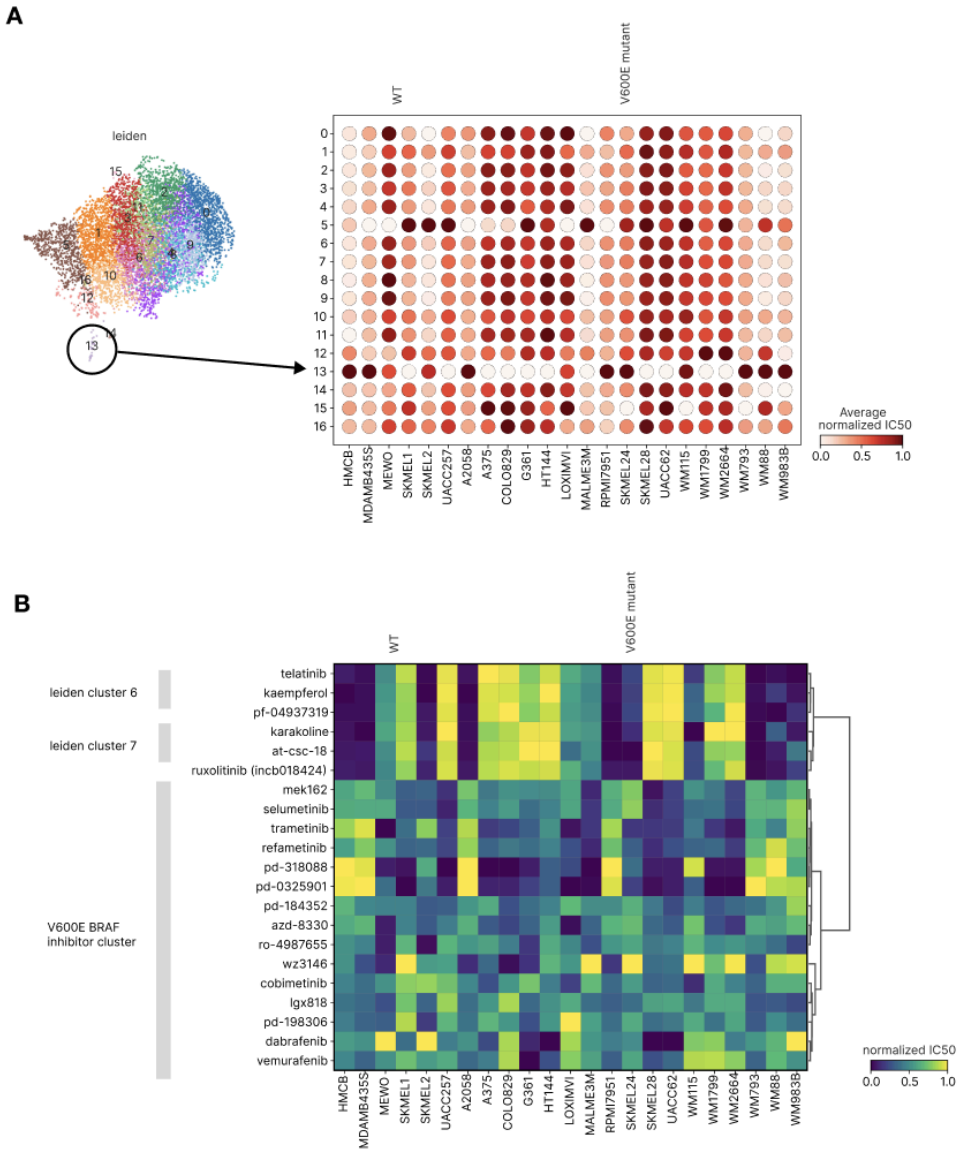


**Fig. S8. In silico modeling of perturbation effect in V600E *BRAF* mutant and wild-type cell lines.** A) The space of 23 V600E mutant or wild-type melanoma cell lines, as in Fig. 3d, colored by leiden clustering. The dotplot depicts the average $IC_{50}$ value across compounds per cluster, with darker clusters therefore indicating more resistance to the compounds in that cluster. The left 6 cell lines are putatively wild-type and remaining cell lines are putatively V600E *BRAF* mutants. B) Prophet predictions for all compounds from leiden cluster 13 as well as 3 random compounds selected from leiden clusters 6 and 7 for contrast. Prophet predicts a distinctly different pattern of behavior for compounds in cluster 13, which also contain known V600E *BRAF* inhibitors.

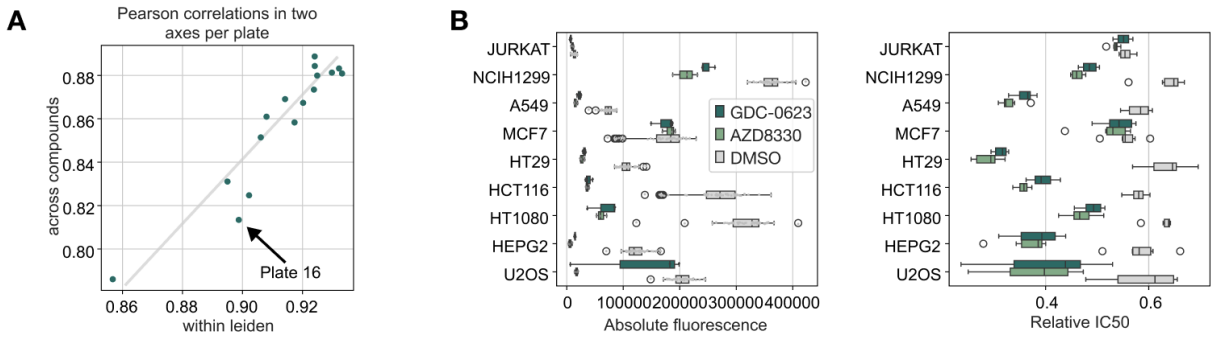


**Fig. S9. Experimental design and in vitro validation.** A) Metrics used to select one out of 16 possible plates in an in vitro drug screening setup. First, Prophet was used to predict the outcome of each 352-compound plate in 9 cell lines. X-axis reports the average pairwise Pearson correlation between compounds within the same leiden cluster. Y-axis reports the average pairwise correlation between all compounds. We optimized for low correlation across compounds and high correlation within cluster (distance from the regression line). B) The same plots as for Fig. 4d, but including all cell lines. Prophet also correctly predicts that AZD8330 has, for most cell lines, a stronger effect on viability than GDC-0623.

## Supplementary Tables

Supplementary Table 1

| Name | Treatment | readout type | # of read outs | # of cell states | # of treatments | # of combinations | # of experiments | Sparsity | Included in Prophet |
|---|---|---|---|---|---|---|---|---|---|
| **Horlbeck et al.**(*27*) | CRISPR combination | viability | 1 | 2 | 367 | 67384 | 126037 | 0.935 | Yes |
| **GDSC**(*24*) | Drug singleton | IC50 | 1 | 682 | 224 | 224 | 135106 | 0.884 | Yes |
| **GDSC²**(*25*) | Drug combination | IC50 | 1 | 110 | 61 | 1195 | 61795 | 0.470 | Yes |
| **cpg-0000 jump**(*8*) | Drug/ CRISPR singleton | cellpainting features | 200 | 2 | 404 | 404 | 161600 | 1.000 | No |
| **LINCS**(*23*) | Drug singleton | RNA (fluorescence) | 42 | 57 | 7306 | 7306 | 1491002 | 0.085 | Yes |
| **PRISM**(*1*) | Drug singleton | Log2FC | 1 | 556 | 2065 | 2065 | 1121973 | 0.977 | Yes |
| **SCORE**(*3*) | CRISPR singleton | essentiality (CERES) | 1 | 249 | 15573 | 15573 | 1279146 | 0.330 | Yes |
| **Saunders et al.**(*28*) | CRISPR singleton/ combination | celltype proportions | 1 | 620 | 24 | 24 | 9370 | 0.630 | No |
| **CTRP**(*26*) | drug singleton | IC50 | 1 | 822 | 545 | 545 | 357865 | 0.799 | Yes |

**Table S1. Dataset overview.** Shown are the metrics of the datasets employed. The treatments are divided into CRISPR or drugs and combinations or singletons. The readouts range from Viability to $IC_{50}$ or cell type proportions. In the case of JUMP and LINCS, the number of readouts is larger than 1 since there are several cell painting features predicted and a total of 50 selected genes. The number of combinations represents the number of unique combinations of interventions. It is different from the number of interventions just in the case of datasets with combinations of interventions. The number of experiments represents the total number of datapoints in the dataset. Notice that it is not the number of combinations times the number of cell states since some datasets do not measure the entire experimental matrix. The sparsity column denotes how completely the experimental matrix was filled in for seen conditions.

Supplementary Table 2

| Hyperparameter | Value |
|---|---|
| Warmup steps | 10000 |
| Batch size | [2048, 1024, 512, 256] |
| Precision | float32 |
| Embedding dimensionality | [512, 256, 128] |
| Transformer blocks | [4,2] |
| Attention heads | [1,2,4] |

**Table S2. Overview of architecture hyperparameters used for the Prophet-individual models.** Shown are the model hyperparameters used for Prophet-individual models.

Supplementary Table 3

| Hyperparameter | Value |
|---|---|
| | |
| Optimizer | AdamW |
| Optimizer momentum | $\beta_1 = 0.9$, $\beta_2 = 0.99$ |
| Maximum learning rate | 1e-4 |
| Minimum learning rate | 0 |
| Weight decay | 0.01 |
| Dropout | 0.10 |
| Warmup steps | 10000 |
| Batch size | 4096 |
| Precision | float32 |
| Embedding dimensionality | 512 |
| Gene embedding dimensionality | 800 |
| Drug embedding dimensionality | 1219 |
| Cell state embedding dimensionality | 300 |
| Activation function | GeLU |
| Transformer blocks | 8 |
| Attention heads | 8 |
| Number of parameters | 20.1 M |

**Table S3. Overview of pretraining architecture hyperparameters used for the final Prophet pretraining model.** Shown are the model hyperparameters used for Prophet's pretraining.

Supplementary Table 4

| **Pearson correlation** | **No replicates aggregated** | **Intervention replicates aggregated** | **Cell line replicates no aggregated** | **Intervention replicates no aggregated** | **Cell line replicates aggregated** | **All replicates aggregated** |
|---|---|---|---|---|---|---|
| **Between Random replicates** | 0.481 | 0.674 | | - | | - |
| **Within cell line** | 0.615 | 0.820 | | - | | - |
| **Across cell lines** | - | - | | 0.625 | | 0.755 |

**Table S4. Statistical metrics of the SCORE dataset across the cell line axis.** Shown are the Pearson correlation scores obtained on the SCORE dataset evaluating the replicates along the cell line axis. The Pearson correlations are computed between random replicates, between the replicates that belong to a cell line (within a cell line) and across cell lines once the replicates are all aggregated (across cell lines). In each case, they are computed for different scenarios: no replicate is aggregated, just the interventional replicates are aggregated, just the cell line replicates are aggregated, and all replicates are aggregated. Importantly, the Pearson correlation between random replicates once the interventional replicates are aggregated (0.674) is bigger than the Pearson correlation between replicates of a same cell line (0.615). This means that the cell line-specific information gets lost once the replicates are aggregated and, hence, one can not distinguish a random replicate from a real one.

Supplementary Table 5

| **Pearson correlation** | **No replicates aggregated** | **Intervention replicates aggregated** | **Cell line replicates no aggregated** | **Intervention replicates no aggregated** | **Cell line replicates aggregated** | **All replicates aggregated** |
|---|---|---|---|---|---|---|
| **Between random replicates** | 0.029 | - | | 0.046 | | - |
| **Within intervention** | 0.125 | - | | 0.208 | | - |
| **Across interventions** | - | - | | 0.106 | | 0.131 |

**Table S5. Statistical metrics of the SCORE dataset across the intervention axis.** Shown are the Pearson correlation scores obtained on the SCORE dataset evaluating the replicates along the intervention axis. The Pearson correlations are computed between random replicates, between the replicates that belong to an intervention (within intervention) and across interventions once the replicates are all aggregated (across interventions). In each case, they are computed for different scenarios: no replicate is aggregated, just the interventional replicates are aggregated, just the cell line replicates are aggregated, and all replicates are aggregated. Importantly, the Pearson correlation between random replicates once the cell line replicates are aggregated (0.046) is lower than the Pearson correlation between replicates of a same intervention (0.125). This means that the intervention-specific information does not get lost once the replicates are aggregated and, hence, one can distinguish a random replicate from a real one. Therefore, the mean estimator can be outperformed since the experiment retains biological signals.

Supplementary Table 6

| Primary mechanism of action | frac of cluster | n in cluster/all | p-value |
|---|---|---|---|
| AKT inhibitor | .36 | .27 | 0 |
| MEK inhibitor | .55 | .61 | 0 |
| EGFR inhibitor | .71 | .52 | 0 |
| mTOR inhibitor | .33 | .19 | 0 |

**Table S6. Enrichment of known mechanisms of action in four clusters in small molecule $IC_{50}$ space.** Each of the four distinct clusters shown in Fig. 3b describe a unique cell proliferation pathway. To demonstrate that the mechanism of action enrichment is significant, we report three statistics on the annotations in each of cluster: 1) the fraction of the small molecules in the cluster which are known to have the indicated mechanism of action 2) the fraction of all small molecules with that annotation which are in the cluster 3) p-value (Fisher exact test, Supplementary Methods). Values which were smaller than $10^{-7}$ are reported as 0.

Supplementary Table 7

| Compound | Pertpy annotation | Prophet annotation | Correct, if not, target? | Citation (for those not pertpy annotated) |
|---|---|---|---|---|
| gsk2110183b | | PI3K/AKT | TRUE | (*84*) |
| gsk2110183 | AKT1 | PI3K/AKT | TRUE | |
| gdc-0068 | AKT1 | PI3K/AKT | TRUE | |
| gdc-0941 | PIK3CA | PI3K/AKT | TRUE | |
| afuresertib | | PI3K/AKT | TRUE | PubChem |
| azd5363_gdsc | | PI3K/AKT | TRUE | (*85*) |
| ipatasertib | | PI3K/AKT | TRUE | PubChem |
| pictilisib | | PI3K/AKT | PI3K | (*86*) |
| taselisib | PIK3CA | PI3K/AKT | TRUE | |
| uprosertib | AKT1 | PI3K/AKT | TRUE | |
| mk-2206 | AKT1 | PI3K/AKT | TRUE | |
| xl-647 | | EGFR | TRUE | (*87*) |
| pelitinib | EGFR | EGFR | TRUE | |
| poziotinib | EGFR | EGFR | TRUE | |
| pd-153035 | EGFR | EGFR | TRUE | |
| av-412 | EGFR | EGFR | TRUE | |
| cudc-101 | EGFR | EGFR | TRUE | |
| tyrphostin-ag-1478 | EGFR | EGFR | TRUE | |
| azd8931 | EGFR | EGFR | TRUE | |
| neratinib | EGFR | EGFR | TRUE | |
| ast-1306 | EGFR | EGFR | TRUE | |
| dacomitinib | EGFR | EGFR | TRUE | |
| osi-420 | EGFR | EGFR | TRUE | |
| vandetanib | EGFR | EGFR | TRUE | |
| canertinib | AKT1 | EGFR | TRUE | (*88*) |
| whi-p154 | EGFR | EGFR | TRUE | |
| gw-583340 | EGFR | EGFR | TRUE | |
| zm-306416 | EGFR | EGFR | TRUE | |
| azd3759 | EGFR | EGFR | TRUE | |
| sapitinib | | EGFR | TRUE | (*88*, *89*) |
| ibrutinib | BLK | EGFR | TRUE | (*43*) |

| | | | | |
|---|---|---|---|---|
| erlotinib | EGFR | EGFR | TRUE | |
| osimertinib | EGFR | EGFR | TRUE | |
| gefitinib | EGFR | EGFR | TRUE | |
| lapatinib | EGFR | EGFR | TRUE | |
| erk-inhibitor-11e | | MAPK/MEK | TRUE | (*90*) |
| azd-8330 | | MAPK/MEK | TRUE | DrugBank |
| mek162 | | MAPK/MEK | TRUE | (*91*) |
| pd-184352 | MAP2K1 | MAPK/MEK | TRUE | |
| bvd-523 | | MAPK/MEK | TRUE | (*92*) |
| tak-733 | MAP2K1 | MAPK/MEK | TRUE | |
| cobimetinib | MAP2K1 | MAPK/MEK | TRUE | |
| pd-0325901 | MAP2K1 | MAPK/MEK | TRUE | |
| pd-318088 | | MAPK/MEK | TRUE | (*93*) |
| ro-4987655 | MAP2K1 | MAPK/MEK | TRUE | |
| as-703026 | MAP2K1 | MAPK/MEK | TRUE | |
| pd-198306 | MAP2K1 | MAPK/MEK | TRUE | |
| azd8330 | MAP3K1 | MAPK/MEK | TRUE | |
| vx-11e | MAPK1 | MAPK/MEK | TRUE | |
| ulixertinib | | MAPK/MEK | TRUE | (*93*, *94*) |
| sch772984 | | MAPK/MEK | TRUE | (*95*) |
| pd0325901 | | MAPK/MEK | TRUE | NHI-NCI |
| refametinib | MAP2K1 | MAPK/MEK | TRUE | |
| trametinib | MAP2K1 | MAPK/MEK | TRUE | |
| selumetinib | MAP2K1 | MAPK/MEK | TRUE | |
| azd-8055 | | MTOR | TRUE | (*95*) |
| sb-2343 | MTOR | MTOR | TRUE | |
| ym-201636 | | MTOR | PIKfyve but it's analog PI-103 is | (*96*) |
| azd2014_prism | | MTOR | TRUE | (*97*) |
| mln0128 | MTOR | MTOR | TRUE | |
| ku-0063794 | MTOR | MTOR | TRUE | |
| gne-317 | | MTOR | TRUE | (*98*) |
| azd8055 | MTOR | MTOR | TRUE | |
| tanespimycin | HSP90AA1 | MTOR | HSP90AA1 | DrugBank |

**Table S7. Compounds in four viability-critical pathways.** Compounds which were not automatically annotated by pertpy were manually annotated. The appropriate citation is referenced in the final column.

Supplementary Table 8

| Compound | Other names | Known V600E inhibitor | Clinically approved | Citation re: activity |
|---|---|---|---|---|
| PD-0325901 | Mirdametinib | --- | Phase II, neurofibroma | (*99*) |
| PD-184352 | | --- | No | (*100*) |
| Dabrafenib | | TRUE | Yes | |
| PLX-4720 | | TRUE | No | (*101*) |
| Vemurafenib | PLX-4032 | TRUE | Yes, V600E BRAF melanoma | (*102*) |
| SCH772984 | | TRUE, also for NRAS | No | (*103*) |
| AS-703026 | Pimasertib | TRUE | Phase II, NRAS melanoma | Only in combination with vermurafenib (*104*) |
| BVD-523 | Ulixertinib | TRUE | Phase II, GI cancer | (*105*, *106*) |
| Cobimetinib | | TRUE | Yes | |
| MEK162 | Binimetinib | TRUE | Yes, for V600E BRAF cancer | |
| Trametinib | | TRUE | Yes, V600E BRAF melanoma | |
| MK-0752 | | Known that is isn't selective | Phase I, various non-melanoma cancers | (*107*) |
| VX-11E | | Known that it is broadly toxic | No | (*106*) |
| TAK-733 | | Known that it isn't selective | Phase I | (*108*) |
| Selumetinib | | True, but it is NOT a V600E melanoma inhibitor | Yes, for neurofibroma | (*109*) |
| WZ3146 | | FALSE | No | No main reference |
| AZD-8330 | | FALSE | Phase I, solid tumor | (*110*) |
| PD-198306 | | FALSE | No | |
| PD-318088 | | FALSE | Preclinical at Pfizer | (*110*) |
| Refametinib | | FALSE | Phase II, hepato carcinoma | (*111*) |
| RO-4987655 | | FALSE | Phase I | (*112*) |

**Table S8. Compounds in a cluster containing V600E *BRAF* melanoma treatments.**
Additional information citing each compound as indicated for either melanoma or as a cancer

treatment is referenced in the final column. The vast majority of compounds have literature backing as chemotherapies.

Supplementary Table 9

| Cell Line | Gene Mutations | Tissue of Origin |
|---|---|---|
| JURKAT | TP53, NOTCH1, PTEN | T-cell Leukemia (Blood) |
| NCI-H1299 | TP53 | Lung Carcinoma |
| A549 | KRAS, TP53, STK11 | Lung Carcinoma |
| MCF7 | PIK3CA, TP53, ESR1 | Breast Adenocarcinoma |
| HT29 | BRAF, TP53, APC | Colorectal Adenocarcinoma |
| HCT116 | KRAS, TP53, PIK3CA | Colorectal Carcinoma |
| HT1080 | NRAS, TP53 | Fibrosarcoma |
| HEPG2 | TP53, CTNNB1 | Hepatocellular Carcinoma (Liver) |
| U2OS | TP53, RB1 | Osteosarcoma (Bone) |

**Table S9. The nine cell lines available for in vitro validation.** Including known mutation statuses and tissue of origin.

# References


1. S. M. Corsello, R. T. Nagari, R. D. Spangler, J. Rossen, M. Kocak, J. G. Bryan, R. Humeidi, D. Peck, X. Wu, A. A. Tang, V. M. Wang, S. A. Bender, E. Lemire, R. Narayan, P. Montgomery, U. Ben-David, C. W. Garvie, Y. Chen, M. G. Rees, N. J. Lyons, J. M. McFarland, B. T. Wong, L. Wang, N. Dumont, P. J. O'Hearn, E. Stefan, J. G. Doench, C. N. Harrington, H. Greulich, M. Meyerson, F. Vazquez, A. Subramanian, J. A. Roth, J. A. Bittker, J. S. Boehm, C. C. Mader, A. Tsherniak, T. R. Golub, Discovering the anti-cancer potential of non-oncology drugs by systematic viability profiling. *Nat Cancer* **1**, 235–248 (2020).

2. W. Yang, J. Soares, P. Greninger, E. J. Edelman, H. Lightfoot, S. Forbes, N. Bindal, D. Beare, J. A. Smith, I. R. Thompson, S. Ramaswamy, P. A. Futreal, D. A. Haber, M. R. Stratton, C. Benes, U. McDermott, M. J. Garnett, Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **41**, D955–61 (2013).

3. F. M. Behan, F. Iorio, G. Picco, E. Gonçalves, C. M. Beaver, G. Migliardi, R. Santos, Y. Rao, F. Sassi, M. Pinnelli, R. Ansari, S. Harper, D. A. Jackson, R. McRae, R. Pooley, P. Wilkinson, D. van der Meer, D. Dow, C. Buser-Doepner, A. Bertotti, L. Trusolino, E. A. Stronach, J. Saez-Rodriguez, K. Yusa, M. J. Garnett, Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* **568**, 511–516 (2019).

4. C. Bock, P. Datlinger, F. Chardon, M. A. Coelho, M. B. Dong, K. A. Lawson, T. Lu, L. Maroc, T. M. Norman, B. Song, G. Stanley, S. Chen, M. Garnett, W. Li, J. Moffat, L. S. Qi, R. S. Shapiro, J. Shendure, J. S. Weissman, X. Zhuang, High-content CRISPR screening. *Nature Reviews Methods Primers* **2**, 1–23 (2022).

5. M. P. Menden, D. Wang, M. J. Mason, B. Szalai, K. C. Bulusu, Y. Guan, T. Yu, J. Kang, M. Jeon, R. Wolfinger, T. Nguyen, M. Zaslavskiy, AstraZeneca-Sanger Drug Combination DREAM Consortium, I. S. Jang, Z. Ghazoui, M. E. Ahsen, R. Vogel, E. C. Neto, T. Norman, E. K. Y. Tang, M. J. Garnett, G. Y. D. Veroli, S. Fawell, G. Stolovitzky, J. Guinney, J. R. Dry, J. Saez-Rodriguez, Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. *Nat. Commun.* **10**, 2674 (2019).

6. S. N. Chandrasekaran, J. Ackerman, E. Alix, D. Michael Ando, J. Arevalo, M. Bennion, N. Boisseau, A. Borowa, J. D. Boyd, L. Brino, P. J. Byrne, H. Ceulemans, C. Ch'ng, B. A. Cimini, D.-A. Clevert, N. Deflaux, J. G. Doench, T. Dorval, R. Doyonnas, V. Dragone, O. Engkvist, P. W. Faloon, B. Fritchman, F. Fuchs, S. Garg, T. J. Gilbert, D. Glazer, D. Gnutt, A. Goodale, J. Grignard, J. Guenther, Y. Han, Z. Hanifehlou, S. Hariharan, D. Hernandez, S. R. Horman, G. Hormel, M. Huntley, I. Icke, M. Iida, C. B. Jacob, S. Jaensch, J. Khetan, M. Kost-Alimova, T. Krawiec, D. Kuhn, C.-H. Lardeau, A. Lembke, F. Lin, K. D. Little, K. R. Lofstrom, S. Lotfi, D. J. Logan, Y. Luo, F. Madoux, P. A. Marin Zapata, B. A. Marion, G. Martin, N. J. McCarthy, L. Mervin, L. Miller, H. Mohamed, T. Monteverde, E. Mouchet, B. Nicke, A. Ogier, A.-L. Ong, M. Osterland, M. Otrocka, P. J. Peeters, J. Pilling, S. Prechtl, C. Qian, K. Rataj, D. E. Root, S. K. Sakata, S. Scrace, H. Shimizu, D. Simon, P. Sommer, C. Spruiell, I. Sumia, S. E. Swalley, H. Terauchi, A. Thibaudeau, A. Unruh, J. Van de Waeter, M. Van Dyck, C. van Staden, M. Warchoł, E. Weisbart, A. Weiss, N. Wiest-Daessle, G. Williams, S. Yu, B. Zapiec, M. Żyła, S. Singh, A. E. Carpenter, JUMP Cell Painting dataset: morphological impact of 136,000 chemical and genetic perturbations, *bioRxiv* (2023)p. 2023.03.23.534023.

7. S. R. Srivatsan, J. L. McFaline-Figueroa, V. Ramani, L. Saunders, J. Cao, J. Packer, H. A. Pliner, D. L. Jackson, R. M. Daza, L. Christiansen, F. Zhang, F. Steemers, J. Shendure, C. Trapnell, Massively multiplex chemical transcriptomics at single-cell resolution. *Science* **367**, 45–51 (2020).

8. S. N. Chandrasekaran, B. A. Cimini, A. Goodale, L. Miller, M. Kost-Alimova, N. Jamali, J. G. Doench, B. Fritchman, A. Skepner, M. Melanson, A. A. Kalinin, J. Arevalo, M. Haghighi, J. C. Caicedo, D. Kuhn, D. Hernandez, J. Berstler, H. Shafqat-Abbasi, D. E. Root, S. E. Swalley, S. Garg, S. Singh, A. E. Carpenter, Three million images and morphological profiles of cells treated with matched chemical and genetic perturbations. *Nat. Methods* **21**, 1114–1121 (2024).

9. D. Wang, J. Hensman, G. Kutkaite, T. S. Toh, A. Galhoz, GDSC Screening Team, J. R. Dry, J. Saez-Rodriguez, M. J. Garnett, M. P. Menden, F. Dondelinger, A statistical framework for assessing pharmacological responses and biomarkers using uncertainty estimates. *Elife* **9** (2020).

10. D. M. Camacho, K. M. Collins, R. K. Powers, J. C. Costello, J. J. Collins, Next-Generation Machine Learning for Biological Networks. *Cell* **173**, 1581–1592 (2018).

11. A. Partin, T. S. Brettin, Y. Zhu, O. Narykov, A. Clyde, J. Overbeek, R. L. Stevens, Deep learning methods for drug response prediction in cancer: Predominant and emerging trends. *Front. Med.* **10**, 1086097 (2023).

12. C. Pacini, J. M. Dempster, I. Boyle, E. Gonçalves, H. Najgebauer, E. Karakoc, D. van der Meer, A. Barthorpe, H. Lightfoot, P. Jaaks, J. M. McFarland, M. J. Garnett, A. Tsherniak, F. Iorio, Integrated cross-study datasets of genetic dependencies in cancer. *Nat. Commun.* **12**, 1661 (2021).

13. D. Yao, L. Binan, J. Bezney, B. Simonton, J. Freedman, C. J. Frangieh, K. Dey, K. Geiger-Schuller, B. Eraslan, A. Gusev, A. Regev, B. Cleary, Scalable genetic screening for regulatory circuits using compressed Perturb-seq. *Nat. Biotechnol.*, doi: 10.1038/s41587-023-01964-9 (2023).

14. T. Chu, T. T. Nguyen, B. D. Hai, Q. H. Nguyen, T. Nguyen, Graph Transformer for Drug Response Prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **20**, 1065–1072 (2023).

15. W. Peng, T. Chen, W. Dai, Predicting Drug Response Based on Multi-Omics Fusion and Graph Convolution. *IEEE J Biomed Health Inform* **26**, 1384–1393 (2022).

16. Y.-C. Tang, A. Gottlieb, Explainable drug sensitivity prediction through cancer pathway enrichment. *Sci. Rep.* **11**, 3128 (2021).

17. Y. Zhu, Z. Ouyang, W. Chen, R. Feng, D. Z. Chen, J. Cao, J. Wu, TGSA: protein-protein association-based twin graph neural networks for drug response prediction with similarity augmentation. *Bioinformatics* **38**, 461–468 (2022).

18. Y. L. Chow, S. Singh, A. E. Carpenter, G. P. Way, Predicting drug polypharmacology from cell morphology readouts using variational autoencoder latent space arithmetic. *PLoS Comput. Biol.* **18**, e1009888 (2022).

19. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M.

Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, *arXiv [cs.CV]* (2020). http://arxiv.org/abs/2010.11929.

20. T. Hayes, R. Rao, H. Akin, N. J. Sofroniew, D. Oktay, Z. Lin, R. Verkuil, V. Q. Tran, J. Deaton, M. Wiggert, R. Badkundri, I. Shafkat, J. Gong, A. Derry, R. S. Molina, N. Thomas, Y. Khan, C. Mishra, C. Kim, L. J. Bartie, M. Nemeth, P. D. Hsu, T. Sercu, S. Candido, A. Rives, Simulating 500 million years of evolution with a language model, *bioRxiv* (2024)p. 2024.07.01.600583.

21. T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language Models are Few-Shot Learners. *Adv. Neural Inf. Process. Syst.* **abs/2005.14165** (2020).

22. G. P. Way, T. Natoli, A. Adeboye, L. Litichevskiy, A. Yang, X. Lu, J. C. Caicedo, B. A. Cimini, K. Karhohs, D. J. Logan, M. H. Rohban, M. Kost-Alimova, K. Hartland, M. Bornholdt, S. N. Chandrasekaran, M. Haghighi, E. Weisbart, S. Singh, A. Subramanian, A. E. Carpenter, Morphology and gene expression profiling provide complementary information for mapping cell state. *cels* **13**, 911–923.e9 (2022).

23. A. Subramanian, R. Narayan, S. M. Corsello, D. D. Peck, T. E. Natoli, X. Lu, J. Gould, J. F. Davis, A. A. Tubelli, J. K. Asiedu, D. L. Lahr, J. E. Hirschman, Z. Liu, M. Donahue, B. Julian, M. Khan, D. Wadden, I. C. Smith, D. Lam, A. Liberzon, C. Toder, M. Bagul, M. Orzechowski, O. M. Enache, F. Piccioni, S. A. Johnson, N. J. Lyons, A. H. Berger, A. F. Shamji, A. N. Brooks, A. Vrcic, C. Flynn, J. Rosains, D. Y. Takeda, R. Hu, D. Davison, J. Lamb, K. Ardlie, L. Hogstrom, P. Greenside, N. S. Gray, P. A. Clemons, S. Silver, X. Wu, W.-N. Zhao, W. Read-Button, X. Wu, S. J. Haggarty, L. V. Ronco, J. S. Boehm, S. L. Schreiber, J. G. Doench, J. A. Bittker, D. E. Root, B. Wong, T. R. Golub, A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437–1452.e17 (2017).

24. F. Iorio, T. A. Knijnenburg, D. J. Vis, G. R. Bignell, M. P. Menden, M. Schubert, N. Aben, E. Gonçalves, S. Barthorpe, H. Lightfoot, T. Cokelaer, P. Greninger, E. van Dyk, H. Chang, H. de Silva, H. Heyn, X. Deng, R. K. Egan, Q. Liu, T. Mironenko, X. Mitropoulos, L. Richardson, J. Wang, T. Zhang, S. Moran, S. Sayols, M. Soleimani, D. Tamborero, N. Lopez-Bigas, P. Ross-Macdonald, M. Esteller, N. S. Gray, D. A. Haber, M. R. Stratton, C. H. Benes, L. F. A. Wessels, J. Saez-Rodriguez, U. McDermott, M. J. Garnett, A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **166**, 740–754 (2016).

25. P. Jaaks, E. A. Coker, D. J. Vis, O. Edwards, E. F. Carpenter, S. M. Leto, L. Dwane, F. Sassi, H. Lightfoot, S. Barthorpe, D. van der Meer, W. Yang, A. Beck, T. Mironenko, C. Hall, J. Hall, I. Mali, L. Richardson, C. Tolley, J. Morris, F. Thomas, E. Lleshi, N. Aben, C. H. Benes, A. Bertotti, L. Trusolino, L. Wessels, M. J. Garnett, Effective drug combinations in breast, colon and pancreatic cancer cells. *Nature* **603**, 166–173 (2022).

26. A. Basu, N. E. Bodycombe, J. H. Cheah, E. V. Price, K. Liu, G. I. Schaefer, R. Y. Ebright, M. L. Stewart, D. Ito, S. Wang, A. L. Bracha, T. Liefeld, M. Wawer, J. C. Gilbert, A. J. Wilson, N. Stransky, G. V. Kryukov, V. Dancik, J. Barretina, L. A. Garraway, C. S.-Y. Hon, B. Munoz, J. A. Bittker, B. R. Stockwell, D. Khabele, A. M. Stern, P. A. Clemons, A. F. Shamji, S. L. Schreiber, An interactive resource to identify cancer genetic and lineage dependencies

targeted by small molecules. *Cell* **154**, 1151–1161 (2013).

27. M. A. Horlbeck, A. Xu, M. Wang, N. K. Bennett, C. Y. Park, D. Bogdanoff, B. Adamson, E. D. Chow, M. Kampmann, T. R. Peterson, K. Nakamura, M. A. Fischbach, J. S. Weissman, L. A. Gilbert, Mapping the Genetic Landscape of Human Cells. *Cell* **174**, 953–967.e22 (2018).

28. L. M. Saunders, S. R. Srivatsan, M. Duran, M. W. Dorrity, B. Ewing, T. H. Linbo, J. Shendure, D. W. Raible, C. B. Moens, D. Kimelman, C. Trapnell, Embryo-scale reverse genetics at single-cell resolution. *Nature* **623**, 782–791 (2023).

29. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention Is All You Need, *arXiv [cs.CL]* (2017). http://arxiv.org/abs/1706.03762.

30. J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *arXiv [cs.CL]* (2018). http://arxiv.org/abs/1810.04805.

31. J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A. A. Margolin, S. Kim, C. J. Wilson, J. Lehár, G. V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M. F. Berger, J. E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jané-Valbuena, F. A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I. H. Engels, J. Cheng, G. K. Yu, J. Yu, P. Aspesi Jr, M. de Silva, K. Jagtap, M. D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R. C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J. P. Mesirov, S. B. Gabriel, G. Getz, K. Ardlie, V. Chan, V. E. Myer, B. L. Weber, J. Porter, M. Warmuth, P. Finan, J. L. Harris, M. Meyerson, T. R. Golub, M. P. Morrissey, W. R. Sellers, R. Schlegel, L. A. Garraway, The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

32. G. Landrum, RDKit: A software suite for cheminformatics, computational chemistry, and predictive modeling. http://www.rdkit.org/RDKit_Overview.pdf.

33. H. L. Morgan, The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *J. Chem. Doc.* **5**, 107–113 (1965).

34. Ž. Avsec, V. Agarwal, D. Visentin, J. R. Ledsam, A. Grabska-Barwinska, K. R. Taylor, Y. Assael, J. Jumper, P. Kohli, D. R. Kelley, Effective gene expression prediction from sequence by integrating long-range interactions. *Nat. Methods* **18**, 1196–1203 (2021).

35. M.-A. Bray, S. Singh, H. Han, C. T. Davis, B. Borgeson, C. Hartland, M. Kost-Alimova, S. M. Gustafsdottir, C. C. Gibson, A. E. Carpenter, Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat. Protoc.* **11**, 1757–1774 (2016).

36. A. Kolesnikov, L. Beyer, X. Zhai, J. Puigcerver, J. Yung, S. Gelly, N. Houlsby, Big Transfer (BiT): General Visual Representation Learning, *arXiv [cs.CV]* (2019). http://arxiv.org/abs/1912.11370.

37. X. Zhai, A. Kolesnikov, N. Houlsby, L. Beyer, Scaling Vision Transformers, *arXiv [cs.CV]* (2021). http://arxiv.org/abs/2106.04560.

38. I. C. Henderson, J. C. Allegra, T. Woodcock, S. Wolff, S. Bryan, K. Cartwright, G. Dukart, D.

Henry, Randomized clinical trial comparing mitoxantrone with doxorubicin in previously treated patients with metastatic breast cancer. *J. Clin. Oncol.* **7**, 560–571 (1989).

39. E. J. Feldman, High-dose mitoxantrone in acute leukaemia: New York Medical College experience. *Eur. J. Cancer Care* **6**, 27–32 (1997).

40. S. Vibet, K. Mahéo, J. Goré, P. Dubois, P. Bougnoux, I. Chourpa, Differential subcellular distribution of mitoxantrone in relation to chemosensitization in two human breast cancer cell lines. *Drug Metab. Dispos.* **35**, 822–828 (2007).

41. K. Takara, K. Yamamoto, M. Matsubara, T. Minegaki, M. Takahashi, T. Yokoyama, K. Okumura, Effects of α-adrenoceptor antagonists on ABCG2/BCRP-mediated resistance and transport. *PLoS One* **7**, e30697 (2012).

42. P. G. Richardson, C. S. Mitsiades, J. P. Laubach, S. Lonial, A. A. Chanan-Khan, K. C. Anderson, Inhibition of heat shock protein 90 (HSP90) as a therapeutic strategy for the treatment of myeloma and other cancers. *Br. J. Haematol.* **152**, 367–379 (2011).

43. H. Wu, A. Wang, W. Zhang, B. Wang, C. Chen, W. Wang, C. Hu, Z. Ye, Z. Zhao, L. Wang, X. Li, K. Yu, J. Liu, J. Wu, X.-E. Yan, P. Zhao, J. Wang, C. Wang, E. L. Weisberg, N. S. Gray, C.-H. Yun, J. Liu, L. Chen, Q. Liu, Ibrutinib selectively and irreversibly targets EGFR (L858R, Del19) mutant but is moderately resistant to EGFR (T790M) mutant NSCLC Cells. *Oncotarget* **6**, 31313–31322 (2015).

44. A. Wang, X.-E. Yan, H. Wu, W. Wang, C. Hu, C. Chen, Z. Zhao, P. Zhao, X. Li, L. Wang, B. Wang, Z. Ye, J. Wang, C. Wang, W. Zhang, N. S. Gray, E. L. Weisberg, L. Chen, J. Liu, C.-H. Yun, Q. Liu, Ibrutinib targets mutant-EGFR kinase with a distinct binding conformation. *Oncotarget* **7**, 69760–69769 (2016).

45. K. T. Flaherty, I. Puzanov, K. B. Kim, A. Ribas, G. A. McArthur, J. A. Sosman, P. J. O'Dwyer, R. J. Lee, J. F. Grippo, K. Nolop, P. B. Chapman, Inhibition of mutated, activated BRAF in metastatic melanoma. *N. Engl. J. Med.* **363**, 809–819 (2010).

46. G. S. Falchook, G. V. Long, R. Kurzrock, K. B. Kim, T. H. Arkenau, M. P. Brown, O. Hamid, J. R. Infante, M. Millward, A. C. Pavlick, S. J. O'Day, S. C. Blackman, C. M. Curtis, P. Lebowitz, B. Ma, D. Ouellet, R. F. Kefford, Dabrafenib in patients with melanoma, untreated brain metastases, and other solid tumours: a phase 1 dose-escalation trial. *Lancet* **379**, 1893–1901 (2012).

47. R. B. Corcoran, C. E. Atreya, G. S. Falchook, E. L. Kwak, D. P. Ryan, J. C. Bendell, O. Hamid, W. A. Messersmith, A. Daud, R. Kurzrock, M. Pierobon, P. Sun, E. Cunningham, S. Little, K. Orford, M. Motwani, Y. Bai, K. Patel, A. P. Venook, S. Kopetz, Combined BRAF and MEK Inhibition With Dabrafenib and Trametinib in BRAF V600-Mutant Colorectal Cancer. *J. Clin. Oncol.* **33**, 4023–4031 (2015).

48. A. Hauschild, P. A. Ascierto, D. Schadendorf, J. J. Grob, A. Ribas, F. Kiecker, C. Dutriaux, L. V. Demidov, C. Lebbé, P. Rutkowski, C. U. Blank, R. Gutzmer, M. Millward, R. Kefford, T. Haas, A. D'Amelio Jr, E. Gasal, B. Mookerjee, P. B. Chapman, Long-term outcomes in patients with BRAF V600-mutant metastatic melanoma receiving dabrafenib monotherapy: Analysis from phase 2 and 3 clinical trials. *Eur. J. Cancer* **125**, 114–120 (2020).

49. G. V. Long, K. T. Flaherty, D. Stroyakovskiy, H. Gogas, E. Levchenko, F. de Braud, J.

Larkin, C. Garbe, T. Jouary, A. Hauschild, V. Chiarion-Sileni, C. Lebbe, M. Mandalá, M. Millward, A. Arance, I. Bondarenko, J. B. A. G. Haanen, J. Hansson, J. Utikal, V. Ferraresi, P. Mohr, V. Probachai, D. Schadendorf, P. Nathan, C. Robert, A. Ribas, M. A. Davies, S. R. Lane, J. J. Legos, B. Mookerjee, J.-J. Grob, Dabrafenib plus trametinib versus dabrafenib monotherapy in patients with metastatic BRAF V600E/K-mutant melanoma: long-term survival and safety analysis of a phase 3 study. *Ann. Oncol.* **30**, 1848 (2019).

50. Z. Eroglu, A. Ribas, Combination therapy with BRAF and MEK inhibitors for melanoma: latest evidence and place in therapy. *Ther. Adv. Med. Oncol.* **8**, 48–56 (2016).

51. F. Xing, Y. Persaud, C. A. Pratilas, B. S. Taylor, M. Janakiraman, Q.-B. She, H. Gallardo, C. Liu, T. Merghoub, B. Hefter, I. Dolgalev, A. Viale, A. Heguy, E. De Stanchina, D. Cobrinik, G. Bollag, J. Wolchok, A. Houghton, D. B. Solit, Concurrent loss of the PTEN and RB1 tumor suppressors attenuates RAF dependence in melanomas harboring (V600E)BRAF. *Oncogene* **31**, 446–457 (2012).

52. A. Maklad, A. Sharma, I. Azimi, Calcium Signaling in Brain Cancers: Roles and Therapeutic Targeting. *Cancers* **11** (2019).

53. Y. Zhang, N. Cruickshanks, F. Yuan, B. Wang, M. Pahuski, J. Wulfkuhle, I. Gallagher, A. F. Koeppel, S. Hatef, C. Papanicolas, J. Lee, E. E. Bar, D. Schiff, S. D. Turner, E. F. Petricoin, L. S. Gray, R. Abounader, Targetable T-type Calcium Channels Drive Glioblastoma. *Cancer Res.* **77**, 3479–3490 (2017).

54. M. Sedeeq, A. Maklad, T. Dutta, Z. Feng, R. Wilson, N. Gueven, I. Azimi, T-Type Calcium Channel Inhibitors Induce Apoptosis in Medulloblastoma Cells Associated with Altered Metabolic Activity. *Mol. Neurobiol.* **59**, 2932–2945 (2022).

55. E. Topkas, N. Cai, A. Cumming, M. Hazar-Rethinam, O. M. Gannon, M. Burgess, N. A. Saunders, L. Endo-Munoz, Auranofin is a potent suppressor of osteosarcoma metastasis. *Oncotarget* **7**, 831–844 (2016).

56. Lynch Thomas J., Bell Daphne W., Sordella Raffaella, Gurubhagavatula Sarada, Okimoto Ross A., Brannigan Brian W., Harris Patricia L., Haserlat Sara M., Supko Jeffrey G., Haluska Frank G., Louis David N., Christiani David C., Settleman Jeff, Haber Daniel A., Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib. *N. Engl. J. Med.* **350**, 2129–2139.

57. T. Okabe, I. Okamoto, K. Tamura, M. Terashima, T. Yoshida, T. Satoh, M. Takada, M. Fukuoka, K. Nakagawa, Differential constitutive activation of the epidermal growth factor receptor in non-small cell lung cancer cells bearing EGFR gene mutation and amplification. *Cancer Res.* **67**, 2046–2053 (2007).

58. A. Bairoch, The Cellosaurus, a Cell-Line Knowledge Resource. *J. Biomol. Tech.* **29**, 25–38 (2018).

59. F. Morgillo, T. Cascone, E. D'Aiuto, E. Martinelli, T. Troiani, P. Saintigny, R. De Palma, J. V. Heymach, L. Berrino, C. Tuccillo, F. Ciardiello, Antitumour efficacy of MEK inhibitors in human lung cancer cells and their derivatives with acquired resistance to different tyrosine kinase inhibitors. *Br. J. Cancer* **105**, 382–392 (2011).

60. X. Wen, Y.-C. Ou, G. Bogatcheva, G. Thomas, A. Mahadevan-Jansen, B. Singh, E. C. Lin,

R. Bardhan, Probing metabolic alterations in breast cancer in response to molecular inhibitors with Raman spectroscopy and validated with mass spectrometry. *Chem. Sci.* **11**, 9863–9874 (2020).

61. Y. Gu, M. Helenius, K. Väänänen, D. Bulanova, J. Saarela, A. Sokolenko, J. Martens, E. Imyanitov, S. Kuznetsov, BRCA1-deficient breast cancer cell lines are resistant to MEK inhibitors and show distinct sensitivities to 6-thioguanine. *Sci. Rep.* **6**, 28217 (2016).

62. A. H. Reis, S. Y. Sokol, Rspo2 antagonizes FGF signaling during vertebrate mesoderm formation and patterning. *Development* **147** (2020).

63. A. H. Reis, S. Y. Sokol, Rspo2 inhibits TCF3 phosphorylation to antagonize Wnt signaling during vertebrate anteroposterior axis specification. *Sci. Rep.* **11**, 13433 (2021).

64. C. E. Friedman, Q. Nguyen, S. W. Lukowski, A. Helfer, H. S. Chiu, J. Miklas, S. Levy, S. Suo, J.-D. J. Han, P. Osteil, G. Peng, N. Jing, G. J. Baillie, A. Senabouth, A. N. Christ, T. J. Bruxner, C. E. Murry, E. S. Wong, J. Ding, Y. Wang, J. Hudson, H. Ruohola-Baker, Z. Bar-Joseph, P. P. L. Tam, J. E. Powell, N. J. Palpant, Single-Cell Transcriptomic Analysis of Cardiac Differentiation from Human PSCs Reveals HOPX-Dependent Cardiomyocyte Maturation. *Cell Stem Cell* **23**, 586–598.e8 (2018).

65. A. Manfrin, Y. Tabata, E. R. Paquet, A. R. Vuaridel, F. R. Rivest, F. Naef, M. P. Lutolf, Engineered signaling centers for the spatially controlled patterning of human pluripotent stem cells. *Nat. Methods* **16**, 640–648 (2019).

66. S. M. Morgani, J. J. Metzger, J. Nichols, E. D. Siggia, A.-K. Hadjantonakis, Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate patterning. *Elife* **7** (2018).

67. R. C. Lindsley, J. G. Gill, M. Kyba, T. L. Murphy, K. M. Murphy, Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. *Development* **133**, 3787–3796 (2006).

68. I. Loshchilov, F. Hutter, Decoupled Weight Decay Regularization, *arXiv [cs.LG]* (2017). http://arxiv.org/abs/1711.05101.

69. I. Loshchilov, F. Hutter, Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608. 03983* (2016).

70. X. Glorot, Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks” in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Y. W. Teh, M. Titterington, Eds. (PMLR, Chia Laguna Resort, Sardinia, Italy, 13--15 May 2010)vol. 9 of *Proceedings of Machine Learning Research*, pp. 249–256.

71. D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, C. von Mering, STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607 (2019).

72. T. M. Norman, M. A. Horlbeck, J. M. Replogle, A. Y. Ge, A. Xu, M. Jost, L. A. Gilbert, J. S. Weissman, Exploring genetic interaction manifolds constructed from rich single-cell phenotypes. *Science* **365** (2019).

73. J. M. Replogle, R. A. Saunders, A. N. Pogson, J. A. Hussmann, A. Lenail, A. Guna, L. Mascibroda, E. J. Wagner, K. Adelman, G. Lithwick-Yanai, N. Iremadze, F. Oberstrass, D. Lipson, J. L. Bonnar, M. Jost, T. M. Norman, J. S. Weissman, Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. *Cell* **185** (2022).

74. J. Du, P. Jia, Y. Dai, C. Tao, Z. Zhao, D. Zhi, Gene2vec: distributed representation of genes based on co-expression. *BMC Genomics* **20**, 7–15 (2019).

75. S. A. Aleksander, J. Balhoff, S. Carbon, J. M. Cherry, H. J. Drabkin, D. Ebert, M. Feuermann, P. Gaudet, N. L. Harris, D. P. Hill, R. Lee, H. Mi, S. Moxon, C. J. Mungall, A. Muruganugan, T. Mushayahama, P. W. Sternberg, P. D. Thomas, K. Van Auken, J. Ramsey, D. A. Siegele, R. L. Chisholm, P. Fey, M. C. Aspromonte, M. V. Nugnes, F. Quaglia, S. Tosatto, M. Giglio, S. Nadendla, G. Antonazzo, H. Attrill, G. Dos Santos, S. Marygold, V. Strelets, C. J. Tabone, J. Thurmond, P. Zhou, S. H. Ahmed, P. Asanitthong, B. D. Luna, M. N. Erdol, M. C. Gage, K. M. Ali, K. Y. C. Li, M. Long, A. Michalak, A. Pesala, A. Pritazahra, S. C. C. Saverimuttu, R. Su, K. E. Thurlow, R. C. Lovering, C. Logie, S. Oliferenko, J. Blake, K. Christie, L. Corbani, M. E. Dolan, H. J. Drabkin, D. P. Hill, L. Ni, D. Sitnikov, C. Smith, A. Cuzick, J. Seager, L. Cooper, J. Elser, P. Jaiswal, P. Gupta, P. Jaiswal, S. Naithani, M. Lera-Ramirez, K. Rutherford, V. Wood, J. L. De Pons, M. R. Dwinell, G. T. Hayman, M. L. Kaldunski, A. E. Kwitek, S. J. F. Laulederkind, M. A. Tutaj, M. Vedi, S. J. Wang, P. D'Eustachio, L. Aimo, K. Axelsen, A. Bridge, N. Hyka-Nouspikel, A. Morgat, S. A. Aleksander, J. M. Cherry, S. R. Engel, K. Karra, S. R. Miyasato, R. S. Nash, M. S. Skrzypek, S. Weng, E. D. Wong, E. Bakker, T. Z. Berardini, L. Reiser, A. Auchincloss, K. Axelsen, G. Argoud-Puy, M. C. Blatter, E. Boutet, L. Breuza, A. Bridge, C. Casals-Casas, E. Coudert, A. Estreicher, F. M. Livia, M. Feuermann, A. Gos, N. Gruaz-Gumowski, C. Hulo, N. Hyka-Nouspikel, F. Jungo, P. Le Mercier, D. Lieberherr, P. Masson, A. Morgat, I. Pedruzzi, L. Pourcel, S. Poux, C. Rivoire, S. Sundaram, A. Bateman, E. Bowler-Barnett, H. Bye-A-Jee, P. Denny, A. Ignatchenko, R. Ishtiaq, A. Lock, Y. Lussi, M. Magrane, M. J. Martin, S. Orchard, P. Raposo, E. Speretta, N. Tyagi, K. Warner, R. Zaru, A. D. Diehl, R. Lee, J. Chan, S. Diamantakis, D. Raciti, M. Zarowiecki, M. Fisher, C. James-Zorn, V. Ponferrada, A. Zorn, S. Ramachandran, L. Ruzicka, M. Westerfield, The Gene Ontology knowledgebase in 2023. *Genetics* **224** (2023).

76. M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

77. CZI Single-Cell Biology Program, S. Abdulla, B. Aevermann, P. Assis, S. Badajoz, S. M. Bell, E. Bezzi, B. Cakir, J. Chaffer, S. Chambers, J. Michael Cherry, T. Chi, J. Chien, L. Dorman, P. Garcia-Nieto, N. Gloria, M. Hastie, D. Hegeman, J. Hilton, T. Huang, A. Infeld, A.-M. Istrate, I. Jelic, K. Katsuya, Y. J. Kim, K. Liang, M. Lin, M. Lombardo, B. Marshall, B. Martin, F. McDade, C. Megill, N. Patel, A. Predeus, B. Raymor, B. Robatmili, D. Rogers, E. Rutherford, D. Sadgat, A. Shin, C. Small, T. Smith, P. Sridharan, A. Tarashansky, N. Tavares, H. Thomas, A. Tolopko, M. Urisko, J. Yan, G. Yeretssian, J. Zamanian, A. Mani, J. Cool, A. Carr, CZ CELL×GENE Discover: A single-cell data platform for scalable exploration, analysis and modeling of aggregated data, *bioRxiv* (2023)p. 2023.10.30.563174.

78. D. van der Meer, S. Barthorpe, W. Yang, H. Lightfoot, C. Hall, J. Gilbert, H. E. Francies, M. J. Garnett, Cell Model Passports—a hub for clinical, genetic and functional datasets of

preclinical cancer models. *Nucleic Acids Res.* **47**, D923 (2019).

79. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

80. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, *arXiv [stat.ML]* (2018). http://arxiv.org/abs/1802.03426.

81. A. Tsherniak, F. Vazquez, P. G. Montgomery, B. A. Weir, G. Kryukov, G. S. Cowley, S. Gill, W. F. Harrington, S. Pantel, J. M. Krill-Burger, R. M. Meyers, L. Ali, A. Goodale, Y. Lee, G. Jiang, J. Hsiao, W. F. J. Gerath, S. Howell, E. Merkel, M. Ghandi, L. A. Garraway, D. E. Root, T. R. Golub, J. S. Boehm, W. C. Hahn, Defining a Cancer Dependency Map. *Cell* **170**, 564–576.e16 (2017).

82. L. Heumos, Y. Ji, L. May, T. Green, X. Zhang, X. Wu, J. Ostner, S. Peidli, A. Schumacher, K. Hrovatin, M. Mueller, F. Chong, G. Sturm, A. Tejada, E. Dann, M. Dong, M. Bahrami, I. Gold, S. Rybakov, A. Namsaraeva, A. A. Moinfar, Z. Zheng, E. Roellin, I. Mekki, C. Sander, M. Lotfollahi, H. B. Schiller, F. J. Theis, Pertpy: an end-to-end framework for perturbation analysis, *bioRxiv* (2024)p. 2024.08.04.606516.

83. M. Ester, H. P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD* (1996).

84. M. Dumble, M.-C. Crouthamel, S.-Y. Zhang, M. Schaber, D. Levy, K. Robell, Q. Liu, D. J. Figueroa, E. A. Minthorn, M. A. Seefeld, M. B. Rouse, S. K. Rabindran, D. A. Heerding, R. Kumar, Discovery of Novel AKT Inhibitors with Enhanced Anti-Tumor Effects in Combination with the MEK Inhibitor. *PLoS One* **9** (2014).

85. B. R. Davies, H. Greenwood, P. Dudley, C. Crafter, D. H. Yu, J. Zhang, J. Li, B. Gao, Q. Ji, J. Maynard, S. A. Ricketts, D. Cross, S. Cosulich, C. C. Chresta, K. Page, J. Yates, C. Lane, R. Watson, R. Luke, D. Ogilvie, M. Pass, Preclinical pharmacology of AZD5363, an inhibitor of AKT: pharmacodynamics, antitumor activity, and correlation of monotherapy activity with genetic background. *Mol. Cancer Ther.* **11** (2012).

86. D. Sarker, J. E. Ang, R. Baird, R. Kristeleit, K. Shah, V. Moreno, P. A. Clarke, F. I. Raynaud, G. Levy, J. A. Ware, K. Mazina, R. Lin, J. Wu, J. Fredrickson, J. M. Spoerke, M. R. Lackner, Y. Yan, L. S. Friedman, S. B. Kaye, M. K. Derynck, P. Workman, J. S. de Bono, First-in-human Phase I study of Pictilisib (GDC-0941), a potent pan-class I phosphatidylinositol-3-kinase (PI3K) inhibitor, in patients with advanced solid tumors. *Clin. Cancer Res.* **21**, 77 (2015).

87. M. C. Pietanza, T. J. Lynch, P. N. Lara, J. Cho, R. H. Yanagihara, N. Vrindavanam, N. M. Chowhan, S. M. Gadgeel, N. A. Pennell, R. Funke, B. Mitchell, H. A. Wakelee, V. A. Miller, XL647--a multitargeted tyrosine kinase inhibitor: results of a phase II study in subjects with non-small cell lung cancer who have progressed after responding to treatment with either gefitinib or erlotinib. *J. Thorac. Oncol.* **7** (2012).

88. H. A. M. Gomaa, A. T. Ali, M. Abdel Gabbar, M. A. Kandeil, The Effect of Canertinib on Sensitivity of Cytotoxic Drugs in Tamoxifen-Resistant Breast Cancer Cells In Vitro. *Int. J. Genomics Proteomics* **2018** (2018).

89. Recent advancements of 4-aminoquinazoline derivatives as kinase inhibitors and their

applications in medicinal chemistry. *Eur. J. Med. Chem.* **170**, 55–72 (2019).

90. E. Jasek-Gajda, H. Jurkowska, M. Jasińska, J. A. Litwin, G. J. Lis, Combination of ERK2 inhibitor VX-11e and voreloxin synergistically enhances anti-proliferative and pro-apoptotic effects in leukemia cells. *Apoptosis* **24**, 849 (2019).

91. A. Maiti, K. Naqvi, T. M. Kadia, G. Borthakur, K. Takahashi, P. Bose, N. G. Daver, A. Patel, Y. Alvarado, M. Ohanian, C. D. DiNardo, J. E. Cortes, E. J. Jabbour, G. Garcia-Manero, H. M. Kantarjian, F. Ravandi, Phase II Trial of MEK Inhibitor Binimetinib (MEK162) in RAS-mutant Acute Myeloid Leukemia. *Clin. Lymphoma Myeloma Leuk.* **19** (2019).

92. R. J. Sullivan, J. R. Infante, F. Janku, D. J. L. Wong, J. A. Sosman, V. Keedy, M. R. Patel, G. I. Shapiro, J. W. Mier, A. W. Tolcher, A. Wang-Gillam, M. Sznol, K. Flaherty, E. Buchbinder, R. D. Carvajal, A. M. Varghese, M. E. Lacouture, A. Ribas, S. P. Patel, G. A. DeCrescenzo, C. M. Emery, A. L. Groover, S. Saha, M. Varterasian, D. J. Welsch, D. M. Hyman, B. T. Li, First-in-Class ERK1/2 Inhibitor Ulixertinib (BVD-523) in Patients with MAPK Mutant Advanced Solid Tumors: Results of a Phase I Dose-Escalation and Expansion Study. *Cancer Discov.* **8** (2018).

93. J. F. Ohren, H. Chen, A. Pavlovsky, C. Whitehead, E. Zhang, P. Kuffa, C. Yan, P. McConnell, C. Spessard, C. Banotai, W. T. Mueller, A. Delaney, C. Omer, J. Sebolt-Leopold, D. T. Dudley, I. K. Leung, C. Flamme, J. Warmus, M. Kaufman, S. Barrett, H. Tecle, C. A. Hasemann, Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nat. Struct. Mol. Biol.* **11**, 1192–1197 (2004).

94. M. E. Burkard, M. McKean, J. R. Ahnert, N. B. Mettu, J. C. Jones, J. G. Misleh, W. W. Ma, K.-H. Lim, E. Gabriela Chiorean, M. J. Pishvaian, S. M. Gadgeel, H. A. McKean, B. Kreider, D. Knoerzer, A. Groover, M. L. Varterasian, J. A. Box, C. Emery, R. J. Sullivan, A two-part, phase II, multi-center study of the ERK inhibitor ulixertinib (BVD-523) for patients with advanced malignancies harboring MEK or atypical BRAF alterations (BVD-523-ABC). *J. Clin. Oncol.*, doi: 10.1200/JCO.2022.40.16_suppl.TPS3172 (2022).

95. A. Chaikuad, E. Tacconi, J. Zimmer, Y. Liang, N. S. Gray, M. Tarsounas, S. Knapp, A unique inhibitor binding site in ERK1/2 is associated with slow binding kinetics. *Nat. Chem. Biol.* **10**, 853 (2014).

96. C. Du, X. Guan, J. Yan, Two-pore channel blockade by phosphoinositide kinase inhibitors YM201636 and PI-103 determined by a histidine residue near pore-entrance. *Communications biology* **5** (2022).

97. S. M. Guichard, J. Curwen, T. Bihani, C. M. D'Cruz, J. W. Yates, M. Grondine, Z. Howard, B. R. Davies, G. Bigley, T. Klinowska, K. G. Pike, M. Pass, C. M. Chresta, U. M. Polanska, R. McEwen, O. Delpuech, S. Green, S. C. Cosulich, AZD2014, an Inhibitor of mTORC1 and mTORC2, Is Highly Effective in ER+ Breast Cancer When Administered Using Intermittent or Continuous Schedules. *Mol. Cancer Ther.* **14** (2015).

98. L. Salphati, T. P. Heffron, B. Alicke, M. Nishimura, K. Barck, R. A. Carano, J. Cheong, K. A. Edgar, J. Greve, S. Kharbanda, H. Koeppen, S. Lau, L. B. Lee, J. Pang, E. G. Plise, J. L. Pokorny, H. B. Reslan, J. N. Sarkaria, J. J. Wallin, X. Zhang, S. E. Gould, A. G. Olivero, H. S. Phillips, Targeting the PI3K pathway in the brain--efficacy of a PI3K inhibitor optimized to cross the blood-brain barrier. *Clin. Cancer Res.* **18** (2012).

99. L. Ciuffreda, D. Del Bufalo, M. Desideri, C. Di Sanza, A. Stoppacciaro, M. R. Ricciardi, S. Chiaretti, S. Tavolaro, B. Benassi, A. Bellacosa, R. Foà, A. Tafuri, F. Cognetti, A. Anichini, G. Zupi, M. Milella, Growth-inhibitory and antiangiogenic activity of the MEK inhibitor PD0325901 in malignant melanoma with or without BRAF mutations. *Neoplasia* **11**, 720–731 (2009).

100. V. Roulstone, M. Pedersen, J. Kyula, D. Mansfield, A. A. Khan, G. McEntee, M. Wilkinson, E. Karapanagiotou, M. Coffey, R. Marais, A. Jebar, F. Errington-Mais, A. Melcher, R. Vile, H. Pandha, M. McLaughlin, K. J. Harrington, BRAF- and MEK-Targeted Small Molecule Inhibitors Exert Enhanced Antimelanoma Effects in Combination With Oncolytic Reovirus Through ER Stress. *Mol. Ther.* **23**, 931 (2015).

101. J. Tsai, J. T. Lee, W. Wang, J. Zhang, H. Cho, S. Mamo, R. Bremer, S. Gillette, J. Kong, N. K. Haass, K. Sproesser, L. Li, K. S. M. Smalley, D. Fong, Y.-L. Zhu, A. Marimuthu, H. Nguyen, B. Lam, J. Liu, I. Cheung, J. Rice, Y. Suzuki, C. Luu, C. Settachatgul, R. Shellooe, J. Cantwell, S.-H. Kim, J. Schlessinger, K. Y. J. Zhang, B. L. West, B. Powell, G. Habets, C. Zhang, P. N. Ibrahim, P. Hirth, D. R. Artis, M. Herlyn, G. Bollag, Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proceedings of the National Academy of Sciences* **105**, 3041–3046 (2008).

102. J. T. Lee, L. Li, P. A. Brafford, M. van den Eijnden, M. B. Halloran, K. Sproesser, N. K. Haass, K. S. M. Smalley, J. Tsai, G. Bollag, M. Herlyn, PLX4032, a Potent Inhibitor of the B-Raf V600E Oncogene, Selectively Inhibits V600E-positive Melanomas. *Pigment Cell Melanoma Res.* **23**, 820 (2010).

103. E. J. Morris, S. Jha, C. R. Restaino, P. Dayananth, H. Zhu, A. Cooper, D. Carr, Y. Deng, W. Jin, S. Black, B. Long, J. Liu, E. Dinunzio, W. Windsor, R. Zhang, S. Zhao, M. H. Angagaw, E. M. Pinheiro, J. Desai, L. Xiao, G. Shipps, A. Hruza, J. Wang, J. Kelly, S. Paliwal, X. Gao, B. S. Babu, L. Zhu, P. Daublain, L. Zhang, B. A. Lutterbach, M. R. Pelletier, U. Philippar, P. Siliphaivanh, D. Witter, P. Kirschmeier, W. R. Bishop, D. Hicklin, D. G. Gilliland, L. Jayaraman, L. Zawel, S. Fawell, A. A. Samatar, Discovery of a novel ERK inhibitor with activity in models of acquired resistance to BRAF and MEK inhibitors. *Cancer Discov.* **3**, 742–750 (2013).

104. S. J. Park, S. W. Hong, J. H. Moon, D. H. Jin, J. S. Kim, C. K. Lee, K. P. Kim, Y. S. Hong, E. K. Choi, J. S. Lee, J. L. Lee, T. W. Kim, The MEK1/2 inhibitor AS703026 circumvents resistance to the BRAF inhibitor PLX4032 in human malignant melanoma cells. *Am. J. Med. Sci.* **346** (2013).

105. U. A. Germann, B. F. Furey, W. Markland, R. R. Hoover, A. M. Aronov, J. J. Roix, M. Hale, D. M. Boucher, D. A. Sorrell, G. Martinez-Botella, M. Fitzgibbon, P. Shapiro, M. J. Wick, R. Samadani, K. Meshaw, A. Groover, G. DeCrescenzo, M. Namchuk, C. M. Emery, S. Saha, D. J. Welsch, Targeting the MAPK Signaling Pathway in Cancer: Promising Preclinical Activity with the Novel Selective ERK1/2 Inhibitor BVD-523 (Ulixertinib). *Mol. Cancer Ther.* **16** (2017).

106. T. D. Lebedev, E. R. Khabusheva, S. R. Mareeva, K. A. Ivanenko, A. V. Morozov, P. V. Spirin, P. M. Rubtsov, A. V. Snezhkina, A. V. Kudryavtseva, M. I. Sorokin, A. A. Buzdin, V. S. Prassolov, Identification of cell type–specific correlations between ERK activity and cell viability upon treatment with ERK1/2 inhibitors. *J. Biol. Chem.* **298** (2022).

107. S. Su, G. Chhabra, M. A. Ndiaye, C. K. Singh, T. Ye, W. Huang, C. N. Dewey, V. Setaluri,

N. Ahmad, PLK1 and NOTCH Positively Correlate in Melanoma and Their Combined Inhibition Results in Synergistic Modulations of Key Melanoma Pathways. *Mol. Cancer Ther.* **20**, 161–172 (2021).

108. L. N. Micel, J. J. Tentler, A. C. Tan, H. M. Selby, K. L. Brunkow, K. M. Robertson, S. L. Davis, P. J. Klauck, T. M. Pitts, E. Gangolli, R. Fabrey, S. M. O'Connell, P. W. Vincent, S. G. Eckhardt, Antitumor activity of the MEK inhibitor TAK-733 against melanoma cell lines and patient-derived tumor explants. *Mol. Cancer Ther.* **14** (2015).

109. M. J. Sale, K. Balmanno, J. Saxena, E. Ozono, K. Wojdyla, R. E. McIntyre, R. Gilley, A. Woroniuk, K. D. Howarth, G. Hughes, J. R. Dry, M. J. Arends, P. Caro, D. Oxley, S. Ashton, D. J. Adams, J. Saez-Rodriguez, P. D. Smith, S. J. Cook, MEK1/2 inhibitor withdrawal reverses acquired resistance driven by BRAFV600E amplification whereas KRASG13D amplification promotes EMT-chemoresistance. *Nat. Commun.* **10**, 1–22 (2019).

110. Y. Cheng, H. Tian, Current Development Status of MEK Inhibitors. *Molecules : A Journal of Synthetic Chemistry and Natural Product Chemistry* **22** (2017).

111. H. Y. Lim, J. Heo, H. J. Choi, C.-Y. Lin, J.-H. Yoon, C. Hsu, K.-M. Rau, R. T. P. Poon, W. Yeo, J.-W. Park, M. H. Tay, W.-S. Hsieh, C. Kappeler, P. Rajagopalan, H. Krissel, M. Jeffers, C.-J. Yen, W. Y. Tak, A Phase II Study of the Efficacy and Safety of the Combination Therapy of the MEK Inhibitor Refametinib (BAY 86-9766) Plus Sorafenib for Asian Patients with Unresectable Hepatocellular Carcinoma. *Clin. Cancer Res.* **20**, 5976–5985 (2014).

112. T. Tegnebratt, E. Ruge, S. Bader, N. Ishii, S. Aida, Y. Yoshimura, C.-H. Ooi, L. Lu, N. Mitsios, V. Meresse, J. Mulder, M. Pawlak, M. Venturi, J. Tessier, S. Stone-Elander, Evaluation of efficacy of a new MEK inhibitor, RO4987655, in human tumor xenografts by [18F] FDG-PET imaging combined with proteomic approaches. *EJNMMI Res.* **4** (2014).