Text S1: Prediction Algorithm and Statistical Tests

Overview of the approach

Our approach to identifying basal states and ligand responses predictive of sensitivity to small molecule drugs relies on partial least squares regression (PLSR) followed by a "variable selection" step: the selection of the most informative variables on the basis of the value of the variable importance projection (VIP) (62). PLSR is less prone to overfitting than standard linear algorithms, especially in the case of multi-collinearity (the situation in which variables used to construct a model are highly correlated) because PLSR relies on principal directions rather than individual variables. The use of variable selection ensures that models are easy to interpret by removing many different measurements having low statistical weights.

To validate individual PLSR models and evaluate prediction quality (as measured by q^2 values) we performed leave-one-out cross-validation. In the context of the current data set, which has a limited number of samples, both leave-one-out and ten-fold cross-validation are commonly accepted methods of validation (63). Here, clustering revealed 3 or 4 clusters and some outliers (Fig. 1 and 2, respectively), which implies a need for a minimum of one cell line per cluster plus an outlier (that is 4 or 5 lines) in the test set for proper stratification. However, the total number cell lines for which we have GI_{50} data for any particular drug (which we derived from the literature) is between 20 and 30, which is insufficient for ten-fold cross validation ($30/10 < 5$). To avoid artifacts from arbitrary partitioning, we used leave-one-out cross-validation. A comparison of predicted and measured GI_{50} values yielded a correlation coefficient and a p-value, on the basis of which the significance of the prediction could be judged; we used $p < 0.05$ as a cutoff. The false discovery rate (FDR) was set at 0.15, but the models discussed in detailed in the text (for ErbB-targeted and PI3K/Akt pathway-targeted inhibitors) remained significant even at FDR=0.05.

We optimized the VIP value cutoff for the variable selection internally to the training set using nested cross-validation (64). Overfitting is a concern for this type of optimization process given the small size of the dataset. However, we obtained similar results in terms of variable coefficients and q^2 using PLSR algorithms with and without variable selection. Moreover, the same variables were generally selected across different iterations of the leave-one-out cross-

validation process. Thus, we believe that the PLSR modeling and VIP selection procedures that we used are robust to overfitting.

Details of the algorithm

For each drug and dataset, one cell line is left out (Fig. S5A). With the remaining cell lines, a model is build using PLSR with variables selected according to their VIP values. The number of principal components and the cutoff for the variable selection is optimized through a nested leave-one-out cross validation (Fig. S5B). The model developed is then used to predict the response of the left out cell line (Fig. S5C). The different steps of the algorithm, assuming *n* cell lines labeled A, B … N, are:

1. Cell line A is left out; the remaining (*n-1*) cell lines are used to build a model.
2. To optimize the model and reduce overfitting, the following nested leave-1-out cross-validation is performed:
    a. Cell line B is left out and a PLSR model is built on the (*n-2*) remaining cell lines using α principal components.
    b. The VIP values from the PLSR model of step 2a are calculated (62).
    c. A reduced model of the (*n-2*) cell lines is built using β principal components and only the variables with a VIP value above a given threshold θ.
    d. The model is used to predict the left out cell line B resulting in a predicted value $B(\alpha, \beta, \theta)$ for cell line B using the α principal component from step a and the β component from step 2c and with θ from step c as a threshold for VIP.
    e. Steps 2a-d are looped for each cell line C, D … N. The prediction quality (measured as the q^2) for the (*n-1*) cell lines can be evaluated as $q^2(\alpha, \beta, \theta)$. Values of α and β are scanned from 1 to 15 and θ is scanned from 0 to 3.5 with 0.05 steps resulting in a landscape of q^2 values.
    f. The optimal parameters based on the nested cross-validation across the (*n-1*) cell lines are defined as the triplet $(\bar{\alpha}, \bar{\beta}, \bar{\theta})$ for which:
        - $\bar{\theta} = \max(\{\theta |\ q^2(\alpha, \beta, \theta)$ is in the top 1% of all measured $q^2\})$
        - $(\bar{\alpha}, \bar{\beta}) = \text{argmax}_{\alpha,\beta}\left(q^2(\alpha, \beta, \bar{\theta})\right)$
3. A model based on cell lines B, C …N and on the parameters derived in step 2f is used to predict the value for cell line A.

4. The steps 1-3 are looped across all cell lines B, C … N.
5. The final q^2 is derived from the predicted values for all cell lines (step 3).

The final models as reported in Figure 3A and Data S3-5 are built with the same optimization scheme (step 2), but including all the cell lines.

To assess statistical significance, we use the *p*-value of the Pearson's correlation between the predicted *GI50* values (through the cross-validation) and their measured values. We also report the Pearson's R in the Data S3-5. Calculating the Pearson's r provides a similar significance to a test based on randomization of the dataset, but can be evaluated analytically and is computationally less demanding than are randomized tests. The correlation approach is also more robust than randomization because, for some drugs, multiple cell lines exhibit the same GI_{50} value (in many cases due to limitations of the drug response assay (20)), which can bias the randomization.

All code is written in MATLAB using the standard embedded functions, such as "simpls" for building the PLSR model.

[INSERT FIGURE S1]

Fig. S1. Experimental design and illustration of the collected datasets. (**A**) A panel of 43 cell lines are characterized by measuring (**B**) their steady state levels, their response to 15 growth factors at three time points using (**C**) ELISA (pAKT and pERK) or (**D**) high throughput microscopy (pAKT, pERK, pJNK and pP38), and (**E**) their response to seven cytokines by high throughput microscopy (pERK, pSTAT1, pSTAT3, and NF-κB localization). Response of these cell lines to 43 different kinase inhibitors has been measured independently (**F**).

[INSERT FIGURE S2]

Fig. S2. Correlation among basal levels of phosphorylated RTKs and intracellular kinases. (**A**) Spearman's correlation between pairs of phosphorylated RTKs and kinases across 39 cell lines. (**B**) Projection of the amount of pIGF1R versus the amount of pErbB2 for all cell lines. (**C**) Spearman's correlation between pairs of selected measures across $HER2^{amp}$ cell lines.

[INSERT FIGURE S3]

Fig. S3. Loading and variance of the principal components of the PCA of the basal profiles. (**A**) Coefficients of the first two principal components of the PCA of the basal levels. These measurements are the most variable across the cell lines and can be used to segregate the different clusters in Fig. 1D. (**B**) Variance captured in the first ten principal components.

[INSERT FIGURE S4]

Fig. S4. Correlation between the ligand responses measured by ELISA and high throughput microscopy. (**A**) Distribution of the Pearson's correlation coefficients r for the different responses. (**B**-**D**) Representative graphs of how well the results obtained by ELISA and high-throughput imaging (HTM) observed among ligand-induced responses correlated: (B) an example of one of the least correlated responses to a ligand in all tested cell lines, (C) an example of an intermediate correlated response, (D) an example of a highly correlated response. Each dot represents the response for a particular cell line.

[INSERT FIGURE S5]

Fig. S5. Description of the leave-one-out PLSR algorithm used for prediction. (**A**) Measurements across cell lines are used to predict drug response measured as *GI50*. (**B**) For each drug, the quality of the prediction of the *GI50* for each cell line is assessed by hiding the measurements for a cell line and building a PLSR model to fit the remaining cell lines (optimization of the model is performed through a nested cross-validation, not shown in the figure). (**C**) The GI50 of the left out cell line is predicted with the model. (**D**) For each cell line, the error of the prediction is recorded to evaluate the q^2 (measure of the quality of the prediction) and the p-value (significance of the model). (**E**) A final model is generated and (**F**) with its corresponding fit. (**G**) This process is repeated for each of the 43 drugs.

[INSERT FIGURE S6]

Fig. S6. Correlation between predictions made with ligand responses measured by ELISA or imaging. (**A**) Distribution of the q^2 for drugs for which significant predictions were obtained with the signaling profile measured by ELISA or high-throughput microscopy (HTM). Non-significant predictions (NS) are set to the value zero. Colors represent different classes of drugs targeting different groups of proteins: ErbB, ErbB1 or ErbB2; other RTK, RTKs not in the ErbB family; AKTp, the PI3K/AKT pathway; MAPKp, the MAPK pathway; CC, regulators of the cell cycle; HDAC, Histone deacetylases. (**B**) Coefficients of the variables for drugs targeting the PI3K/Akt pathway that can be significantly predicted with models based on the signaling profile (maximal response across the three measured time points, see table S7). * indicates models based on binarized data that are more predictive than models based on non-binarized data. See data S4 for the values of the model coefficients.

[INSERT FIGURE S7]

Fig. S7. Coefficients of the models based on the signaling profiles for drugs targeting the PI3K/Akt pathway. (A) Coefficients of the variables for the models predictive of drugs targeting the PI3K/Akt pathway and based on the signaling profile obtained using the maximum response across the three time points (see table S7 for dataset description). A dark blue means that the variable is a strong sensitivity marker. * indicates models based on binarized data (i.e. significant responses, see table S7) that are more predictive than models based on non-binarized data. See data S2 for the values of the model coefficients. **(B)** Same as (A) for the models based on the ligand responses measured by ELISA (see data S3). Note that variables on the x-axis are different than in (A).

[INSERT FIGURE S8]

Fig. S8. Distribution of model variables in each cluster for the Lapatinib network representation. Distribution of the model variables for the cell lines that were part of the resistant cluster R1 (in red) or were part of the sensitive clusters S1 (blue) and S2 (purple). For R1, the box plot illustrates the 10-90% range (line), the interquartile (box) and the median (black bar); a collapsed box (e.g. pErbB3) means that more than 75% of the cell lines are below the detection threshold for the given variable; for S1 and S2, the box represents the full range and the line is the median.

Table S1. Names, abbreviations, and UNIPROT ids of the proteins mentioned in the paper.

Protein name	Abbreviation	UNIPROT ID
RAC-alpha serine/threonine-protein kinase	Akt	P31749
Mitogen-activated protein kinase 3	Erk	P27361
Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase	PTEN	P60484
Receptor tyrosine-protein kinase erbB-2	ErbB2	P04626
RAF proto-oncogene serine/threonine-protein kinase	Raf	P04049
Estrogen receptor	ER	Q16405
Progesterone receptor	PR	P06401
Epidermal growth factor receptor	ErbB1	P00533
Receptor tyrosine-protein kinase erbB-3	ErbB3	P21860
Receptor tyrosine-protein kinase erbB-4	ErbB4	Q15303
Hepatocyte growth factor receptor	c-Met	P08581
Insulin-like growth factor 1 receptor	IGF1R	P08069
Platelet-derived growth factor receptor beta	PDGFR	P05622
Vascular endothelial growth factor receptor 1	VEGFR1	P17948
Vascular endothelial growth factor receptor 2	VEGFR2	P35968
Vascular endothelial growth factor receptor 3	VEGFR3	P35916
Mast/stem cell growth factor receptor Kit	c-Kit	P10721
Insulin receptor	InsR	P06213
Fibroblast growth factor receptor 4	FGFR4	P22455
Interleukin-6 receptor subunit alpha	IL6R	P08887
Signal transducer and activator of transcription 1-alpha/beta	STAT1	P42224
Signal transducer and activator of transcription 3	STAT3	P40763

Table S2. Cell lines and culture conditions.

Cell Line	Subtype	Growth Medium used in this study	ATCC recommendation (if different)	T	CO2
BT-20	TNBC	EMEM + 10% heat Inactivated FBS + 1% P/S		37°C	5%
CAMA-1	HR+	EMEM + 10% heat Inactivated FBS + 1% P/S		37°C	5%
MCF7	HR+	DMEM + 10% heat Inactivated FBS + 1% P/S	EMEM + 0.01 mg/ml bovine insulin +10% FBS	37°C	5%
AU-565	HER2amp	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-38	TNBC	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-70	TNBC	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-202	HER2amp	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1187	TNBC	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1395	TNBC	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1419	HER2amp	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1500	HR+	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1428	HR+	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1569	HER2amp	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1806	TNBC	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1937	TNBC	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-1954	HER2amp	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
ZR-75-1	HR+	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
ZR-75-30	HER2amp	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
SK-BR-3	HER2amp	McCoy's 5a Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
MCF-10A	TNBC	MEBM (Lonza/Clonetics, MEGM, Kit Cat No. CC-3150) 100 ng/ml cholera toxin, Do not filter		37°C	5%
MDA-MB-231	TNBC	Leibovitz's L-15 Medium +10% FBS (no Heat-Inactivated) + 1% P/S		37°C	none
MDA-MB-157	TNBC	Leibovitz's L-15 Medium +10% FBS (no Heat-Inactivated) + 1% P/S		37°C	none

MDA-MB-175VII	HR+	Leibovitz's L-15 Medium +10% FBS (no Heat-Inactivated) + 1% P/S		37°C	none
MDA-MB-453	TNBC	Leibovitz's L-15 Medium +10% FBS (no Heat-Inactivated) + 1% P/S		37°C	none
MDA-MB-468	TNBC	Leibovitz's L-15 Medium +10% FBS (no Heat-Inactivated) + 1% P/S		37°C	none
UACC-893	HER2amp	Leibovitz's L-15 Medium +10% FBS (no Heat-Inactivated) + 1% P/S		37°C	none
184-B5	TNBC	MEBM (Lonza/Clonetics as a kit: MEGM, Kit Cat No. CC-3150) 1 ng/ml cholera toxin, Do not filter		37°C	5%
BT-483	HR+	RPMI-1640 Medium + 20% heat Inactivated FBS + 1% P/S 0.01 mg/ml bovine insulin		37°C	5%
BT-549	TNBC	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S 0.023 IU/ml insulin		37°C	5%
T47D	HR+	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S 0.2 Units/ml bovine insulin		37°C	5%
Hs 578T	TNBC	DMEM + 10% heat Inactivated FBS + 1% P/S 0.01 mg/ml bovine insulin		37°C	5%
MDA-MB-134VI	HR+	Leibovitz's L-15 Medium + 20% Fetal Bovine Serum (no Heat-Inactivated) + 1% P/S		37°C	none
MDA-MB-361	HER2amp	Leibovitz's L-15 Medium + 20% Fetal Bovine Serum (no Heat-Inactivated) + 1% P/S		37°C	none
BT-474	HER2amp	RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S	ATCC Hybri-Care Medium + 10% FBS + 1.5 g/L sodium bicarbonate	37°C	5%
UACC-812	HER2amp	Leibovitz's L-15 medium + 20% FBS (no Heat-Inactivated) + 1% P/S 2 mM L-glutamine 20 ng/ml human EGF		37°C	none
MCF 10F	TNBC	1:1 mixture of DMEM and Ham's F12 (+ L-glutamine + 15mM Hepes, Gibco, Cat# 11330) 20 ng/ml Human epidermal growth factor 100 ng/ml cholera toxin 0.01 mg/ml bovine insulin 500 ng/ml hydrocortisone, 95% 5% horse serum		37°C	5%
MCF-12A	TNBC	1:1 mixture of DMEM and Ham's F12 (+ L-glutamine + 15mM Hepes, Gibco, Cat# 11330)		37°C	5%

		20 ng/ml Human epidermal growth factor 100 ng/ml cholera toxin 0.01 mg/ml bovine insulin 500 ng/ml hydrocortisone, 95% 5% horse serum			
MDA-MB-436	TNBC	Leibovitz's L-15 medium + 10% FBS (no Heat-Inactivated) + 1% P/S 10 mg/l insulin	Leibovitz's L-15 medium + 10% FBS (no Heat-Inactivated) + 1% P/S 10 mg/l insulin 16 mg/l glutathione, 90%	37°C	none
MDA-MB-415	HR+	Leibovitz's L-15 medium + 15% FBS (no Heat-Inactivated) + 1% P/S 2mM L-glutamine 10 mg/l insulin	Leibovitz's L-15 medium + 15% FBS (no Heat-Inactivated) + 1% P/S 2mM L-glutamine 10 mg/l insulin 10 mg/l glutathione, 85%	37°C	none
DU-4475		RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-2218		RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%
HCC-2157		RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S 0.02mg/ml Insulin 50nM Hydrocortisone 1ng/ml EGF		37°C	5%
HCC-1599		RPMI-1640 Medium + 10% heat Inactivated FBS + 1% P/S		37°C	5%

Table S3. Description of the ELISA kits used for measuring the basal profiles.

Receptor Family	Abbreviation	capture antibody vendor and label	epitope	source animal	detection antibody vendor and label	epitope	source animal
-	ERK	Cell Signaling Technologies 7050	Total ERK1 Total ERK2	mouse	Cell Signaling Technologies 7050	Total ERK1 Total ERK2	rabbit
-	Akt	Cell Signaling Technologies 7142	Total Akt	rabbit	Cell Signaling Technologies 7142	Total Akt1	mouse
ErbB	EGFR	R&D Systems AF231	Total EGFR	goat	R&D Systems BAF231	Total EGFR	goat
ErbB	ErbB2	R&D Systems MAB1129	Total ErbB2	mouse	R&D Systems BAF1129	Total ErbB2	goat
ErbB	ErbB3	R&D Systems MAB3481	Total ErbB3	mouse	R&D Systems BAM348	Total ErbB3	mouse
ErbB	ErbB4	R&D Systems DYC1133	Total ErbB4	mouse	R&D Systems DYC1133	Total ErbB4	mouse
InsR	InsR	R&D Systems DYC1544	Total Insulin R	mouse	R&D Systems DYC1544	Total Insulin R	mouse
InsR	IGF1R	R&D Systems DYC305	Total IGF1R	mouse	R&D Systems DYC305	Total IGF1R	goat
InsR	IGF2R	R&D Systems DY2447	Total IGF2R	mouse	R&D Systems DY2447	Total IGF2R	goat
VEGFR	VEGFR1	R&D Systems DYC4347	Total VEGFR1	mouse	R&D Systems DYC4347	Total VEGFR1	goat
VEGFR	VEGFR2	R&D Systems DYC1780	Total VEGFR2	mouse	R&D Systems DYC1780	Total VEGFR2	goat
VEGFR	VEGFR3	R&D Systems DYC3491	Total VEGFR3	mouse	R&D Systems DYC3491	Total VEGFR3	mouse
PDGFR	PDGFRa	Cell Signaling Technologies 7264	Total PDGFRa	rabbit	Cell Signaling Technologies 7264	Total PDGFRa	mouse
PDGFR	PDGFRb	R&D Systems DYC385	Total PDGFRb	mouse	R&D Systems DYC385	Total PDGFRb	goat
Met	Met	Invitrogen CHO0285	Total Met	unknown	Invitrogen CHO0285	Total Met	rabbit
ckit	ckit	R&D Systems DY332	Total ckit	mouse	R&D Systems DY332	Total ckit	goat
EphA2	EphA2	R&D Systems DYC3035	Total EphA2	mouse	R&D Systems DYC3035	Total EphA2	goat
Src	Src	R&D Systems 7992	Total Src	rabbit	R&D Systems 7992	Total Src	mouse
TrkA	TrkA	R&D Systems DYC175	Total TrkA	mouse	R&D Systems DYC175	Total TrkA	mouse
FGFR	FGFR1	Merrimack A1	Total FGFR1	human	Merrimack H7	Total FGFR1	human
FGFR	FGFR2	R&D Systems DYC665	Total FGFR2	mouse	R&D Systems DYC665	Total FGFR2	mouse
FGFR	FGFR3	R&D Systems DYC766	Total FGFR3	mouse	R&D Systems DYC766	Total FGFR3	mouse
FGFR	FGFR4	R&D Systems DYC685	Total FGFR4	mouse	R&D Systems DYC685	Total FGFR4	rat
EpCa	EpCAM	R&D Systems DY960	Total EpCAM	mouse	R&D Systems DY960	Total EpCAM	goat

-	ppERK (pErk R&D) *	R&D Systems DYC1018	Total ERK1 Total ERK2	goat	R&D Systems DYC1018	Phospho-ERK1 (Thr202/Tyr204) Phospho-ERK2 (Thr185/Tyr187)	rabbit
-	pERK (pErk CST) **	Cell Signaling Technologies 7246	Phospho-ERK1 (Thr202/Tyr204) Phospho-ERK2 (Thr185/Tyr187)	rabbit	Cell Signaling Technologies 7246	Total ERK1 Total ERK2	mouse
-	pAkt	Millipore 05-591MG	Akt1 PH Domain	mouse	Cell Signaling Technologies 5102	Phospho-Akt (Ser473)	mouse
ErbB	pEGFR	R&D Systems AF231	Total EGFR	goat	Millipore 16-452	Phosphotyrosine	mouse
ErbB	pErbB2	R&D Systems MAB1129	Total ErbB2	mouse	Millipore 16-452	Phosphotyrosine	mouse
ErbB	pErbB3	R&D Systems MAB3481	Total ErbB3	mouse	Millipore 16-452	Phosphotyrosine	mouse
ErbB	pErbB4	R&D Systems DYC2115	Total ErbB4	mouse	Millipore 16-452	Phosphotyrosine	mouse
InsR	pInsR	R&D Systems DYC2718	Total Insulin R	mouse	Millipore 16-452	Phosphotyrosine	mouse
InsR	pIGF1R	R&D Systems DYC1770	Total IGF1R	mouse	Millipore 16-452	Phosphotyrosine	mouse
InsR	pIGF1R-Y1131	Cell Signaling Technologies 7820	Phospho-IGF1R (Tyr1131)	rabbit	Cell Signaling Technologies 7820	Total IGF1R	mouse
VEGFR	pVEGFR1	R&D Systems DYC4170	Total VEGFR1	mouse	Millipore 16-452	Phosphotyrosine	mouse
VEGFR	pVEGFR2	R&D Systems DYC1766	Total VEGFR2	mouse	Millipore 16-452	Phosphotyrosine	mouse
VEGFR	pVEGFR3	R&D Systems DYC2724	Total VEGFR3	mouse	Millipore 16-452	Phosphotyrosine	mouse
PDGFR	pPDGFRa	R&D Systems DYC2114	Total PDGFRa	mouse	Millipore 16-452	Phosphotyrosine	mouse
PDGFR	pPDGFRb	R&D Systems DYC1767	Total PDGFRb	mouse	Millipore 16-452	Phosphotyrosine	mouse
Met	pMet	R&D Systems DYC2480	Total Met	goat	Millipore 16-452	Phosphotyrosine	mouse
ckit	pckit	R&D Systems DYC3527	Total ckit	rat	Millipore 16-452	Phosphotyrosine	mouse
EphA2	pEphA2	R&D Systems DYC4056	Total EphA2	mouse	Millipore 16-452	Phosphotyrosine	mouse
Src	pSrc	R&D Systems DYC2685	Total Src	goat	R&D Systems DYC2685	Phospho-Src (Y419)	rabbit
TrkA	pTrkA	R&D Systems DYC2578	Total TrkA	mouse	Millipore 16-452	Phosphotyrosine	mouse
HR	ER	R&D Systems DYC5715	Total ER alpha	sheep	R&D Systems DYC5715	Total ER alpha	sheep
HR	PR	R&D Systems DYC5415	Total PR	sheep	R&D Systems DYC5415	Total PR	sheep

* assay measures exclusively doubly phosphorylated form of ERK1 and ERK2

** assay measures singly and doubly phosphorylated forms of ERK1 and ERK2

Table S4. Description of ligands used for the measuring the signaling profiles.

Ligand Family	Ligand name	Abbreviation	vendor/catalog no.	UniProt id	Stock concentration
ErbB	Epidermal Growth Factor	EGF	Peprotech AF-100-15	P01133	100µg/ml
ErbB	Epiregulin	EPR	Peprotech 100-04	O14944	100µg/ml
ErbB	Betacellulin	BTC	Peprotech 100-50	P35070	100µg/ml
ErbB	Heregulin β1	HRG	Peprotech 100-03	Q02297	100µg/ml
Ins-IGF	Insulin	INS	Sigma I9278	P01308	100µg/ml
Ins-IGF	Insulin-like Growth Factor 1	IGF-1	Peprotech 100-11	P05019	100µg/ml
Ins-IGF	Insulin-like Growth Factor 2	IGF-2	Peprotech 100-12	P01344	100µg/ml
PDGF	Platelet Derived Growth Factor BB	PDGF-BB	Peprotech 100-14B	P01127	100µg/ml
HGF	Hepatocyte Growth Factor	HGF	Peprotech 100-39	P14210	100µg/ml
SCF	Stem Cell Factor	SCF	Peprotech 300-07	P21583	100µg/ml
FGF	Fibroblast Growth Factor (acidic)	FGF-1	Peprotech 100-17A	P05230	100µg/ml
FGF	Fibroblast Growth Factor (basic)	FGF-2	Peprotech 100-18B	P09038	100µg/ml
NGF	Nerve Growth Factor	NGF-beta	Peprotech 450-01	P01138	100µg/ml
EFNA	Ephrin-A1	EFNA1	R&D 602-A1-200	P20827	100µg/ml
VEGF	Vascular endothelial growth factor A	VEGF165	Peprotech 100-20	P15692	100µg/ml
Interleukin	Interleukin 1	IL-1 alpha	Peprotech 200-01A	P01583	100µg/ml
Interleukin	Interleukin 2	IL-2	Peprotech 200-02	P60568	100µg/ml
Interleukin	Interleukin 6	IL-6	Peprotech 200-06	P05231	100µg/ml
Interferon receptor	Interferon-α *	IFN-alpha	R&D 11200-2	P01562 / P01563	10E6 U/ml stock
Interferon receptor	Interferon-γ	IFN-gamma	Peprotech 300-02	P01579	100µg/ml
Toll-like receptor	Lipopolysaccharide	LPS	Invivogen tlrl-pelps		100µg/ml
TNF receptors	Tumor Necrosis Factor-α	TNF-a	Peprotech 300-01A	P01375	100µg/ml

* Mix of IFN alpha-A and IFN alpha-D

Table S5. Description of the antibodies used for the high-througput microscopy imaging assays. Shading indicates antibodies used in experiments for predicting drug response.

antibody vendor and label	epitope	source animal	secondary antibody	Assay
Cell Signaling Technologies 9106BC	Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204)	mouse	Alexa Fluor 488 Donkey Anti-Mouse IgG (Invitrogen A21202)	Cytokines (CK)
Cell Signaling Technologies 9167BC	Phospho-Stat1 (Tyr701)	rabbit	Alexa Fluor 488 Donkey Anti-Rabbit IgG (H+L) (Invitrogen A21206)	Cytokines (CK)
Cell Signaling Technologies 9145BC	Phospho-Stat3 (Tyr705)	rabbit	Alexa Fluor 488 Donkey Anti-Rabbit IgG (H+L) (Invitrogen A21206)	Cytokines (CK)
Santa Cruz Biotechnology sc-8008	amino acids 1-286 mapping at the N-terminus of NFκB p65	mouse	Alexa Fluor 488 Donkey Anti-Mouse IgG (Invitrogen A21202)	Cytokines (CK)
Cell Signaling Technologies 9106BC	Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204)	mouse	Alexa Fluor 647 Donkey Anti-Mouse IgG (H+L) (Invitrogen A31571)	Cytokines (CK)
Cell Signaling Technologies 9167BC	Phospho-Stat1 (Tyr701)	rabbit	Alexa Fluor 647 Donkey Anti-Rabbit IgG (H+L) (Invitrogen A31573)	Cytokines (CK)
Cell Signaling Technologies 9145BC	Phospho-Stat3 (Tyr705)	rabbit	Alexa Fluor 647 Donkey Anti-Rabbit IgG (H+L) (Invitrogen A31573)	Cytokines (CK)
Santa Cruz Biotechnology sc-8008	amino acids 1-286 mapping at the N-terminus of NFκB p65	mouse	Alexa Fluor 647 Donkey Anti-Mouse IgG (H+L) (Invitrogen A31571)	Cytokines (CK)
Cell Signaling Technologies 9106BC	Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204)	mouse	Alexa Fluor 488 Donkey Anti-Mouse IgG (Invitrogen A21202)	Growth factors (GF)
Cell Signaling Technologies 4051BC	Phospho-Akt (Ser473)	mouse	Alexa Fluor 488 Donkey Anti-Mouse IgG (Invitrogen A21202)	Growth factors (GF)
Cell Signaling Technologies 9216BC	Phospho-p38 MAPK (Thr180/Tyr182)	mouse	Alexa Fluor 488 Donkey Anti-Mouse IgG (Invitrogen A21202)	Growth factors (GF)
Cell Signaling Technologies 9255BC	Phospho-SAPK/JNK (Thr183/Tyr185)	mouse	Alexa Fluor 488 Donkey Anti-Mouse IgG (Invitrogen A21202)	Growth factors (GF)
Cell Signaling Technologies 4370BC	Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204)	rabbit	Alexa Fluor 647 Donkey Anti-Rabbit IgG (H+L) (Invitrogen A31573)	Growth factors (GF)
Cell Signaling Technologies 4060BC	Phospho-Akt (Ser473)	rabbit	Alexa Fluor 647 Donkey Anti-Rabbit IgG (H+L) (Invitrogen A31573)	Growth factors (GF)
Cell Signaling Technologies 4511BC	Phospho-p38 MAPK (Thr180/Tyr182)	rabbit	Alexa Fluor 647 Donkey Anti-Rabbit IgG (H+L) (Invitrogen A31573)	Growth factors (GF)
Cell Signaling Technologies 9251BC	Phospho-SAPK/JNK (Thr183/Tyr185)	rabbit	Alexa Fluor 647 Donkey Anti-Rabbit IgG (H+L) (Invitrogen A31573)	Growth factors (GF)

Table S6. Descriptions of the datasets used for the prediction of *GI50* values. See table S4 for the list of ligands used for the experimental results.

Dataset name	Abbreviation	Measures	Transformation	Size
Basal profile	B	Basal expression and phosphorylation levels measured by ELISA (see **table S3** for protein list).	Each variable is independently normalized between 0 and 1 across all cell lines.	39 cell lines x 46 measured proteins = 39 x 46 values
Binarized Basal profile	Bin-B	Basal expression and phosphorylation levels measured by ELISA (see **table S4** for protein list).	Each variable is independently binarized, 1 corresponding to a value above the median (removing threshold values).	39 cell lines x 46 measured proteins = 39 x 46 values
Signaling profile (growth factors; average across time points)	S(GF)mean	Measures of pERK, pAKT, pP38 and pJNK after growth factor stimulation (see **table S4** for ligand list).	The three time points are averaged. Each variable is independently normalized between 0 and 1 across all cell lines.	38 cell lines x 4 measured proteins x 15 ligands x 2 concentrations = 38 x 120 values
Binarized signaling profile (growth factors; average across time points)	Bin-S(GF)mean	Measures of pERK, pAKT, pP38 and pJNK after growth factor stimulation (see **table S4** for ligand list).	Each measure is set to 1 if the response is 3 standard deviations higher than the control, 0 otherwise. The three time points are then averaged.	38 cell lines x 4 measured proteins x 15 ligands x 2 concentrations = 38 x 120 values
Signaling profile (growth factors; maximal value across time points)	S(GF)max	Measures of pERK, pAKT, pP38 and pJNK after growth factor stimulation (see **table S4** for ligand list).	The value is the maximum of the three time points. Each variable is independently normalized between 0 and 1 across all cell lines.	38 cell lines x 4 measured proteins x 15 ligands x 2 concentrations = 38 x 120 values
Binarized signaling profile (growth factors; maximal value across time points)	Bin- S(GF)max	Measures of pERK, pAKT, pP38 and pJNK after growth factor stimulation (see **table S4** for ligand list).	The value is set to 1 if the response at any of the three time points is 3 standard deviations higher than the control, 0 otherwise.	38 cell lines x 4 measured proteins x 15 ligands x 2 concentrations = 38 x 120 values

Signaling profile (growth factors and cytokines; average across time points)	S(GF+CK)mean	Measures of pERK, pAKT, pP38 and pJNK after growth factor stimulation and measures of pERK, STAT-1, STAT-3 and NF-kB after cytokine stimulation (see **table S4** for ligand list).	The three time points are averaged. Each variable is independently normalized between 0 and 1 across all cell lines.	37 cell lines x 4 measured proteins x (15+7) ligands x 2 concentrations = 37 x 176 values
Binarized signaling profile (growth factors and cytokines; average across time points)	Bin-S(GF+CK)mean	Measures of pERK, pAKT, pP38 and pJNK after growth factor stimulation and measures of pERK, STAT-1, STAT-3 and NF-kB after cytokine stimulation (see **table S4** for ligand list).	Each measure is set to 1 if the response is 3 standard deviations higher than the control, 0 otherwise. The three time points are then averaged.	37 cell lines x 4 measured proteins x (15+7) ligands x 2 concentrations = 37 x 176 values
Basal and signaling profiles (growth factors and cytokines; average across time points)	B+S(GF+CK)mean	Basal expression and phosphorylation levels measured by ELISA; measures of pERK, pAKT, pP38 and pJNK after growth factor stimulation, and measures of pERK, STAT-1, STAT-3 and NF-kB after cytokine stimulation (see **table S4** for ligand list).	For the ligand response, the three time points are averaged. Then, each variable is independently normalized between 0 and 1 across all cell lines.	37 cell lines x 4 measured proteins x (15+7) ligands x 2 concentrations + 37 cell lines x 46 measured proteins = 37 x 222 values
Binarized basal and signaling profiles (growth factors and cytokines; average across time points)	Bin- B+S (GF+CK)mean	Basal expression and phosphorylation levels measured by ELISA; measures of pERK, pAKT, pP38 and pJNK after growth factor stimulation, and measures of pERK, STAT-1, STAT-3 and NF-kB after cytokine stimulation (see **table S4** for ligand list).	For the basal levels, Each variable is independently binarized, 1 corresponding to a value above the median (removing threshold values). For the ligand response, each measure is set to 1 if the response is 3 standard deviations higher than the control, 0 otherwise. Then the three time points are then averaged.	37 cell lines x 4 measured proteins x (15+7) ligands x 2 concentrations + 37 cell lines x 46 measured proteins = 37 x 222 values

Table S7. Enrichment analysis of the GI50 values by *PTEN* and *PI3KCA* mutational status.

Drug name	Median GI50 value for PIK3CA/PTEN WT in log10(M)	Median GI50 value for PIK3CA/PTEN mutant in log10(M)	p-value
AG1478	-4.268	-4.569	0.865
Erlotinib	-4.773	-4.389	0.631
Gefitinib	-4.965	-5.134	0.556
Afatinib	-6.016	-6.199	0.773
Lapatinib	-4.778	-4.781	0.667
Triciribine	-5.622	-5.849	0.648
A6730 SIGMA	-5.490	-5.727	0.427
Rapamycin	-7.331	-6.921	0.580
Temsirolimus	-6.300	-6.252	0.979
BEZ235	-5.969	-6.231	0.291
GSK1059615	-6.263	-6.237	0.835
GSK2119563	-6.018	-6.003	0.819
GSK2126458	-8.012	-8.051	0.884
TGX-221	-4.814	-5.103	0.293
AS-252424	-4.875	-4.701	0.687
GSK1487371	-5.781	-5.580	0.607

Data S1: Measures of the ligand response and basal data grouped by dataset as described in table S6 and raw values. A description of the different spreadsheets can be found in the document.

Data S2: Signaling-targeted drugs used for drug sensitivity predictions and their corresponding reported GI50 in the cell lines.

Data S3: Results of the predictions and coefficients for the significant models for the different datasets used. The description of the datasets used can be found in table S6.

Data S4: Results of the predictions and coefficients for significant models using the ligand responses measured by ELISA.

Data S5: Results of the stratified models using the basal profile data and model coefficients.