

Integrative modeling identifies key determinants of inhibitor sensitivity in breast cancer cell lines

Jastrzebski, Katarzyna; Thijssen, Bram; Kluin, Roelof J.C.; De Lint, Klaas; Majewski, Ian J.; Beijersbergen, Roderick L.; Wessels, Lodewyk F.A.

DOI

[10.1158/0008-5472.CAN-17-2698](https://doi.org/10.1158/0008-5472.CAN-17-2698)

Publication date

2018

Document Version

Final published version

Published in

Cancer Research

Citation (APA)

Jastrzebski, K., Thijssen, B., Kluin, R. J. C., De Lint, K., Majewski, I. J., Beijersbergen, R. L., & Wessels, L. F. A. (2018). Integrative modeling identifies key determinants of inhibitor sensitivity in breast cancer cell lines. *Cancer Research*, 78(15), 4396-4410. <https://doi.org/10.1158/0008-5472.CAN-17-2698>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Integrative Modeling Identifies Key Determinants of Inhibitor Sensitivity in Breast Cancer Cell Lines

Katarzyna Jastrzebski¹, Bram Thijssen^{1,2}, Roelof J.C. Kluin³, Klaas de Lint¹, Ian J. Majewski⁴, Roderick L. Beijersbergen¹, and Lodewyk F.A. Wessels^{1,2,5}

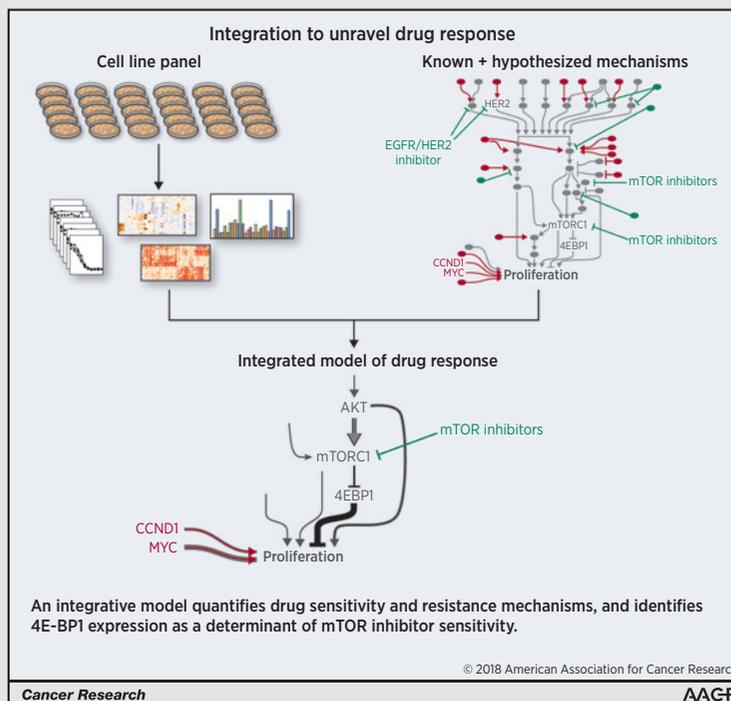


Abstract

Cancer cell lines differ greatly in their sensitivity to anticancer drugs as a result of different oncogenic drivers and drug resistance mechanisms operating in each cell line. Although many of these mechanisms have been discovered, it remains a challenge to understand how they interact to render an individual cell line sensitive or resistant to a particular drug. To better understand this variability, we profiled a panel of 30 breast cancer cell lines in the absence of drugs for their mutations, copy number aberrations, mRNA, protein expression and protein phosphorylation, and for response to seven different kinase inhibitors. We then constructed a knowledge-based, Bayesian computational model that integrates these data types and estimates the relative contribution of various drug sensitivity mechanisms. The resulting model of regulatory signaling explained the majority of the variability observed in drug response. The model also identified cell lines with an unexplained response, and for these we searched for novel explanatory factors. Among others, we found that 4E-BP1 protein expression, and not just the extent of phosphorylation, was a determinant of mTOR inhibitor sensitivity. We validated this finding experimentally and found that overexpression of 4E-BP1 in cell lines that normally possess low levels of this protein is sufficient to increase mTOR inhibitor sensitivity. Taken together, our work demonstrates that combining experimental characterization with integrative modeling can be used to systematically test and extend our understanding of the variability in anticancer drug response.

Significance: By estimating how different oncogenic mutations and drug resistance mechanisms affect the response of cancer cells to kinase inhibitors, we can better understand and ultimately predict response to these anticancer drugs.

Graphical Abstract: <http://cancerres.aacrjournals.org/content/canres/78/15/4396/F1.large.jpg>. *Cancer Res*; 78(15); 4396–410. ©2018 AACR.



¹Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, the Netherlands. ²Oncode Institute, The Netherlands Cancer Institute, Amsterdam, the Netherlands. ³Genomic Sequencing Facility, The Netherlands Cancer Institute, Amsterdam, the Netherlands. ⁴Division of Cancer and Haematology, The Walter and Eliza Hall Institute, Parkville Victoria, Australia. ⁵Faculty of EEMCS, Delft University of Technology, Delft, the Netherlands.

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

K. Jastrzebski and B. Thijssen contributed equally to this article.

Corresponding Authors: Lodewyk F.A. Wessels, The Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam, NH 1066CX, the Netherlands. Phone: 312-0512-7987; Fax: 312-0669-1383; E-mail: l.wessels@nki.nl; and Roderick L. Beijersbergen, r.beijersbergen@nki.nl

doi: 10.1158/0008-5472.CAN-17-2698

©2018 American Association for Cancer Research.

Major Findings

Our study illustrates the power of combining molecular characterization of a panel of 30 breast cancer cell lines with knowledge-based Bayesian computational modeling to better understand the specific signaling pathways involved in regulating the observed variability in anticancer drug response. This approach allowed us to quantify the contribution of diverse driver mutations and resistance mechanisms in determining the response to seven kinase inhibitors. This led us to uncover a novel sensitivity mechanism for mTOR inhibitors, namely the overexpression of 4E-BP1.

Quick Guide to Equations and Assumptions

To integrate the different data types with knowledge of the regulatory signaling pathways, we created an integrative computational model using a modeling framework we call inference of signaling activity. A challenge in constructing such a model is deciding which aspects of cell biology should be included and which can be omitted. To tackle this challenge, we developed the model iteratively using the procedure shown in Fig. 1A. We started out with a small, simple network including only the signaling nodes EGFR, ERK, AKT, and a node depicting proliferation. We then surveyed the literature for signaling events, molecular mechanisms, and recurrent mutations and copy number aberrations associated with breast cancer or known to be involved in determining drug sensitivity (see Supplementary Tables S9 and S10 for all references that were used). We iteratively added more of these relevant mechanisms to the model. Specifically, for every mechanism we created a new model with additional nodes to represent the mutation, amplification, or signaling molecule. At each iteration, we tested the goodness of fit with the posterior predictive distribution (see section on testing the goodness of fit) and used the marginal likelihood to decide whether the newly added mechanisms should be retained. Finally, we stopped the process of model refinement when further additions no longer increased the marginal likelihood or when computation time grew impractically long.

The resulting model is shown in Fig. 1B and includes growth factors, surface receptors, the MAPK and PI3K pathways, mutations, and copy number aberrations, which occur regularly in breast cancer, kinase inhibitors and their targets, and finally the proliferation of the cells. The signaling molecules in the model (the nodes) are linked using activation functions (the arrows), which describe how the signal is propagated between molecules.

The activity of a signaling molecule i in cell line j , $A_{i,j}$, is calculated as follows:

$$A_{i,j}^* = E_{i,j} \cdot \left(b_i + \sum_{k \in \text{parents}_A} (A_i) s_{k,i} \cdot A_{k,j} + \sum_{k \in \text{parents}_M} (A_i) s_{mut,k,i} \cdot M_{k,j} \right) \quad (\text{A})$$

$$A_{i,j} = \max(\min(A_{i,j}^*, 1), 0)$$

The activity $A_{i,j}^*$ is a linear combination of a base activity b_i , the upstream signaling molecules ($A_{k \in \text{parents}(i,j)}$) with signaling strength $s_{k,i}$, and the upstream mutations ($M_{k \in \text{parents}(i,j)}$) with signaling strength $s_{mut,k,i}$ which is then multiplied by the expression of the signaling molecule itself ($E_{i,j}$). The resulting value is clamped between 0 and 1 to give interpretable values that are comparable throughout the network. For a full description, see the Supplementary Materials and Methods.

Fig. 1C shows the structure of the model in template notation for a small part of the network. To illustrate signal propagation between molecules, consider, for example, S6K. The activity of S6K, represented in the figure by the variable *S6K signal*, is a function of the activity of the upstream kinase mTORC1 (*mTORC1 signal*), as well as of the total amount of S6K in the cell (*S6K expression*). The activation function for calculating S6K signal has several parameters, namely the basal activity (*S6K base signal*), the strength of the link between mTORC1 and S6K (*mTORC1 → S6K strength*). Importantly, the parameters, represented by dashed circles, are shared by all cell lines. That is, the values of the parameters are the same for all cell lines. So, although each cell line can have a different amount of mTORC1 activity, a given amount of mTORC1 signal always gives rise to the same amount of input signal to S6K. Note that the model is not intended to be a precise description of all chemical reactions within the network, but rather an abstract representation of relevant regulatory signaling.

We do not explicitly include feedback signaling events in the model. Although feedback signaling is an important aspect of cellular regulatory networks, we find that even without feedback signals the relative viability after drug treatment can still be described well. This is most likely due to the fact that the activity of feedback loops may still be indirectly reflected in the steady-state signal strengths.

The model provides a framework that allows the integration of all data types to infer the parameter values and signaling activities. Some variables are observed directly, for example the presence of a mutation, and in this case the value of the node is set to the observed value directly. Other variables, namely the protein activity, the untreated growth rate, and the viability after drug treatment, are only observed indirectly. For example, the amount of S6K phosphorylation only indirectly reflects S6K activity. In these cases, we add a random variable, $x_{i,j}$, which models the measurement value and which is dependent on the hidden model variables. We then use a likelihood function for this random variable $x_{i,j}$ to infer the parameters. We use a Student t distribution and center the distribution on $x_{i,j}$. The likelihood can be expressed as

$$P(y_{i,j,k} | \theta) = t(y_{i,j,k} | \mu = x_{i,j}(\theta), \sigma = \sigma_i, \nu = 3), \quad (\text{B})$$

$$x_{i,j}(\theta) = \begin{cases} g_i + (1 - g_i) A_{i,j}(\theta) & \text{for RPPA data} \\ r_j(\theta) & \text{for growth data} \\ D_j(\theta) & \text{for drug response data,} \end{cases} \quad (\text{C})$$

where $x_{i,j}$ models the data as described above, $y_{i,j,k}$ is the measurement data, k indexes the biological replicate measurements, $A_{i,j}$ is the signaling activity defined above in Eq. A, r is the untreated growth rate, and D is the relative viability (the variables r and D are defined in the Supplementary Material and Methods). θ represents a vector of all model parameters in Eq. A.

To infer the parameter values and signaling activities, we used Bayesian statistics. For all parameters, prior probability distributions were specified to describe what values the parameters might assume *a priori* (Supplementary Tables S11 and S12).

Subsequently, two variants of Monte Carlo sampling were used to calculate the posterior probability distribution of all parameters (Supplementary Fig. S1).

The major assumptions are thus as follows. First, we assume that the signaling activity is a linear combination of its upstream inputs. Similarly, proliferation is a linear combination of the activity of several effector signaling molecules. Second, we do not explicitly model feedback events, but assume that their effect is indirectly reflected in the steady-state signal strengths. Third, we assume that a signaling molecule with a certain amount of activity always gives rise to the same amount of input activity to its downstream nodes in every cell line, as the signaling strength $s_{k,i}$ is constant across cell lines. However, note that cell lines can still have widely different signaling activity due to variation in the mutations present as well as different gene expression levels.

Introduction

Breast cancer is the most commonly diagnosed cancer in women and the second leading cause of cancer-related death (1). Decades of research have increased our knowledge of the molecular basis of this disease, whereas recent large-scale genomics studies have provided detailed information on mutations, copy number aberrations, and gene expression across different tumor samples, including breast cancer (2, 3). However, despite this increasing amount of knowledge, response rates for cancer treatment remain very low for many cancer subtypes, including breast cancer.

One of the challenges of cancer treatment is the genetic complexity of the disease, involving different oncogenic drivers or combinations thereof that allow the tumor to grow and proliferate. In some subtypes of breast cancer, the major oncogenic drivers are known, as is the case in HER2-amplified breast cancer, associated with the overexpression and aberrant activation of HER2 receptor signaling. Targeted treatment against HER2 indeed provides clinical benefit (4, 5). However, intrinsic resistance is frequently encountered, and acquired resistance often develops in tumors that are initially sensitive (6). A variety of drug resistance mechanisms to HER2-targeted therapies have been discovered in cell lines. For example, activation of the PI3K pathway resulting from mutations in PI3K (7), the loss of PTEN (8), or autocrine HGF signaling (9) have all been reported as mechanisms of drug resistance. However, due to the multitude of resistance mechanisms, which is further complicated by the cross-talk in downstream signaling, it is unclear to what extent each of these mechanisms is important for determining the sensitivity of a particular cell line or tumor. In other types of breast cancer, in particular triple-negative breast cancer, the regulatory signaling that drives the growth of the tumor is even less clear, although the PI3K/AKT pathway deregulation has been identified as a recurrent event (2).

Similar to patients, cell lines show a large degree of variability in drug response (10–12). Understanding the heterogeneous response in cell lines is an important starting point for understanding patient response. But despite many efforts, fully explaining drug response *in vitro* remains a challenge. Computational modeling can be used as a tool capable of untangling the complexity of drug sensitivity and resistance across different cell lines. There have been various approaches to computational modeling of drug sensitivity. These approaches can be broadly divided into two categories: approaches using linear or black box statistical models (11, 13, 14), and approaches using more detailed mechanistic computational models (15, 16). Although both approaches have recovered known drug sensitivity mechanisms and identified several novel associations, they have several limita-

tions. For the black box statistical models, such as elastic net regression (10, 11), random forests (13), support vector machines (13), or a clustering-based method named ACME (14), the models are not sufficiently detailed to capture how molecular characteristics affect drug sensitivity. For example, the interactions between molecular aberrations are not explicitly modeled. This precludes finding all but the strongest associations, despite the very large number of cell lines that were profiled. In addition, available knowledge of signaling pathways is not employed, which could increase the statistical power to find molecular mechanism that associate with drug sensitivity. In the more detailed, mechanistic computational models (15, 16) such background knowledge is used; however, in these cases the number of cell lines studied has been limited, and it is thus unclear to what extent the particular mechanisms are important for explaining the variability across a larger set of cell lines. In addition, these mechanistic modeling studies used only a single data type, for example (phospho)protein expression, limiting the insight into the impact of other molecular aberrations present in the cell lines examined.

To address these limitations, we set out to combine detailed computational modeling of drug sensitivity mechanisms with extensive measurements of multiple data types derived from a breast cancer cell line panel. We developed a combined experimental/computational modeling approach, which can utilize background knowledge from the literature and integrate diverse types of data, including DNA sequencing, RNA sequencing, protein expression, and protein phosphorylation, with drug response data. We subsequently employed the computational model to analyze how the regulatory signaling in each cell line influences response to each drug. The computational model can also be used to identify cases where drug response cannot be explained fully by the existing knowledge. In one case, this lead us to identify and confirm the level of expression of eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1) as a determinant of response to mTOR inhibitors in breast cancer cell lines.

Together, we show the utility of employing integrative computational modeling to combine prior knowledge with measurements of multiple molecular data types to systematically test and extend our understanding of drug response to kinase inhibitors.

Materials and Methods

Detailed materials and methods are provided in the Supplementary Materials and Methods.

Cell line panel

A panel of 30 breast cancer cell lines was assembled from various sources, the details and growth conditions of which are

listed in Supplementary Table S1. Cell lines were authenticated by suppliers using short tandem repeat profiling. The SK-BR-7 cell line was obtained from an internal NKI cell bank and authenticated by STR profiling. We further confirmed the identity of the lines by comparing their mutation profiles (found by DNA sequencing) with those reported in COSMIC (17). Upon receipt, all lines were first expanded and early passage stocks frozen in liquid nitrogen. Lines were kept in culture for no more than 3 months, after which a new cell aliquot was obtained from frozen stock if needed. All cell lines were tested in-house and found to be negative for mycoplasma. Doubling times for each cell line were estimated by fitting exponential curves to confluence measurements obtained using an InCuCyte FLR/ZOOM instrument (Essen Bioscience). The resulting doubling times are provided in Supplementary Table S2.

Drug response assays

Prior to carrying out drug response assays, cell line seeding densities were optimized. Cells seeded as for the cell doubling time experiments were assessed at the 96-hour endpoint for percentage confluence, and incubated with CellTiter-Blue (CTB; Promega) for a measure of metabolic activity. This was to ensure that cell lines did not exceed 90% confluence at assay endpoint and that the CTB signal at this density was not saturated. Seeding densities used for each cell line are listed in Supplementary Table S1.

For drug response assays, cells were seeded at the optimized density and 24 hours later treated with a 10-point 1:3 dilution series of a number of inhibitors using a MicroLab STAR workstation fitted with $8 \times 1,000 \mu\text{L}$ channels and 96-probe head (Hamilton): AZD8055, top dose 3×10^{-5} mol/L; BEZ235, 1×10^{-5} mol/L; GDC0941, 3×10^{-5} mol/L; MK2206, 3×10^{-5} mol/L; PD0325901, 3×10^{-5} mol/L; lapatinib, 3×10^{-5} mol/L; foretinib, 3×10^{-5} mol/L (all from Selleckchem). Each condition, including an untreated negative control and a phenyl arsine oxide (1×10^{-6} mol/L) treated positive control were set up in technical quadruplicate. Following 72-hour incubation, cells were stained with CTB (1:30 dilution) for 4 hours and the signal measured using an Envision spectrophotometer (Perkin Elmer). In the case of the validation experiments with HCC1806 and HCC1937 cell lines expressing 4E-BP1 or GFP constructs, cells were treated in a 9-point 1:3 dilution series of AZD8055, BEZ235, or GDC0941 using a HP D300 Digital Dispenser (Hewlett-Packard), while all other experimental conditions remained the same. Each assay was carried out in biological triplicate. Each replicate of a dose-response experiment was further analyzed by normalization to the negative and positive control (the normalized data are provided in Supplementary Table S3) and fitting to a four-parameter sigmoid function that allowed for the calculation of the IC_{50} (dose at which viability is 50% of the untreated control). The IC_{50} estimates are provided in Supplementary Table S4. For model inference, full dose-response curve data were used.

Long-term drug response assays

HCC1806 parental, GFP- and 4E-BP1-expressing cells were seeded at 600 cells/well, whereas the HCC1937 panel was seeded at 1,200 cells/well, in 96-well plates. Cells were treated, 24 hours after seeding, with a 9-point 1:3 dilution series of AZD8055 (top dose 3.3×10^{-6} mol/L) or BEZ235 (1.1×10^{-6} mol/L) using a HP D300 Digital Dispenser (Hewlett-Packard). Each condition, including an untreated negative control and a phenyl arsine oxide

(1×10^{-6} mol/L) treated positive control, were set up in technical duplicate. Media and drugs were changed every 3 to 4 days over a period of 10 to 11 days of treatment. Cells were then washed with PBS, fixed with 3.7% formaldehyde/PBS, and stained in 0.1% crystal violet solution. Images of dried, stained cells were digitized on a Perfection V750 PRO scanner (Epson).

Molecular characterization

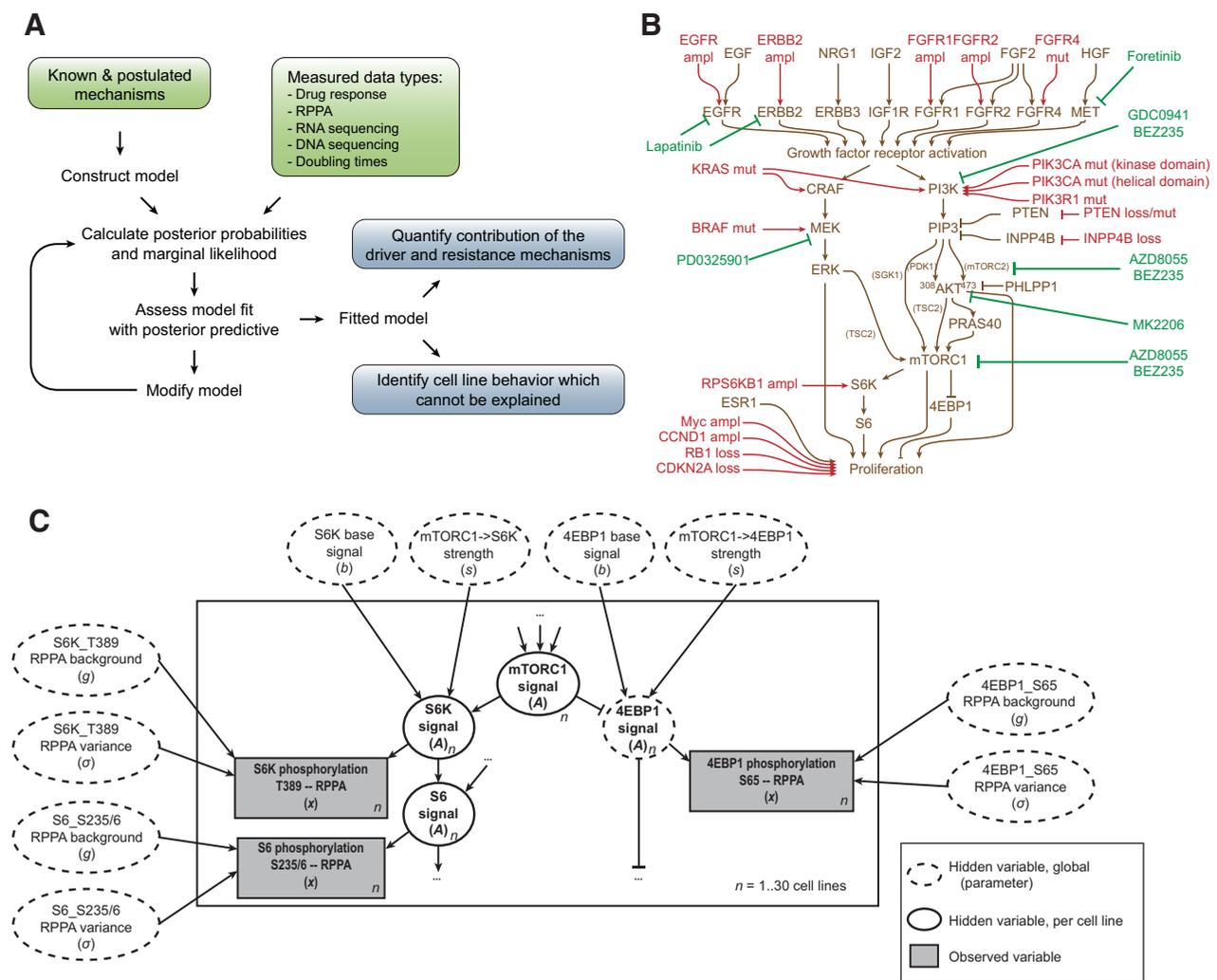
Steady-state RNA and protein expression was determined from cells seeded in 60-mm dishes and grown for 48 hours (seeding densities are listed in Supplementary Table S1). RNA expression was determined using RNA sequencing by the NKI Genomics Facility and protein expression was determined in biological triplicate using reverse-phase protein array (RPPA) analysis by the MD Anderson Cancer Center RPPA Facility. DNA was extracted from cell pellets and targeted sequencing was done using the Human Kinome Capture Kit (Agilent Technologies), containing baits for kinases and cancer-related genes. RNA-sequencing data are available at ArrayExpress, reference E-MTAB-4801, and the normalized read counts are provided in Supplementary Table S5. DNA sequencing data are available at the European Nucleotide Archive, reference PRJEB14120, and the mutation calls and estimated copy number profiles are provided in Supplementary Tables S6 and S7. The RPPA data are included as Supplementary Table S8.

Cap-binding pull-down assays

Cells were seeded in 100-mm dishes (BT549 and CAL-120 at 2.5×10^5 ; Hs 578T at 3×10^5 ; HCC1806 at 4×10^5 ; HCC1937 at 6.25×10^5) and cultured for 48 hours, then treated with AZD8055 (1.11×10^{-7} mol/L), BEZ235 (3.7×10^{-8} mol/L), or vehicle (DMSO) for a further 24 hours. Cells were washed once with ice-cold PBS and lysed in lysis buffer (25 mmol/L Tris-HCl, pH 7.6, 1% Triton X-100, 1 mmol/L DTT) supplemented with cComplete protease and phosSTOP phosphatase inhibitor cocktails (Roche). Lysates were cleared and assayed for protein concentration, then total protein samples were prepared using 20 μg of protein lysate. Cap pull-down samples were prepared by combining 50 μg of total lysate with 20 μL prewashed $m^7\text{GTP}$ -agarose (Jena Bioscience), made up to a total volume of 500 μL with lysis buffer and tumbled at 4°C overnight. The following day, cap pull downs were washed three times in ice-cold lysis buffer, then heated at 70°C for 10 minutes in 20 μL 1 \times Novex LDS Sample Buffer and sample reducing agent. The eluate from the cap pull downs and the total protein control samples were then immediately separated on Novex 4% to 12% gradient gels and immunoblotted using primary antibodies to 4E-BP1, eIF4G, and HSP90 (for total lysate samples only), then reprobed to detect eIF4E protein.

Generation of 4E-BP1-overexpressing cell lines

pLX304-4E-BP1 was obtained from the CCSB-Broad Lentiviral Expression Collection, whereas the pLX304-GFP control construct was generated as outlined previously (18). To produce lentiviral particles, HEK293T cells were cotransfected with the pLX304-4E-BP1 or -GFP bearing construct and a lentiviral packaging mix (pRSV-Rev, pMDLg/pRRE, pCMV-VSV-G; Addgene) using Polyethylenimine (PEI, Linear MW 25,000; Polysciences Inc.). Media were changed 24 hours after transfection. After a further 24 hours, viral supernatant was collected and 0.45- μm filtered. HCC1806 and HCC1937 cells were transduced in the presence of

**Figure 1.**

Modeling procedure and model overviews. **A**, The procedure used to construct the computational models. We used the literature to construct and iteratively update the model until a good fit for the data was obtained. **B**, Simplified overview of the computational model, showing the signaling nodes (brown), the mutations and gene losses and gains (red), and the kinase inhibitors (green). **C**, Detailed view of a part of the computational model as a graphical model in template notation.

hexadimethrine bromide (Sigma-Aldrich) and after 48 hours selected using blasticidin.

Proliferation of 4E-BP1-overexpressing cell lines

HCC1806 parental, GFP-, and 4E-BP1-expressing cells were seeded at 800 cells/well, whereas the HCC1937 panel was seeded at 1,000 cells/well, in 384-well plates with 4 to 6 replicates per condition. Proliferation was monitored using the IncuCyte ZOOM instrument (Essen Biosciences).

Model, data integration, and inference

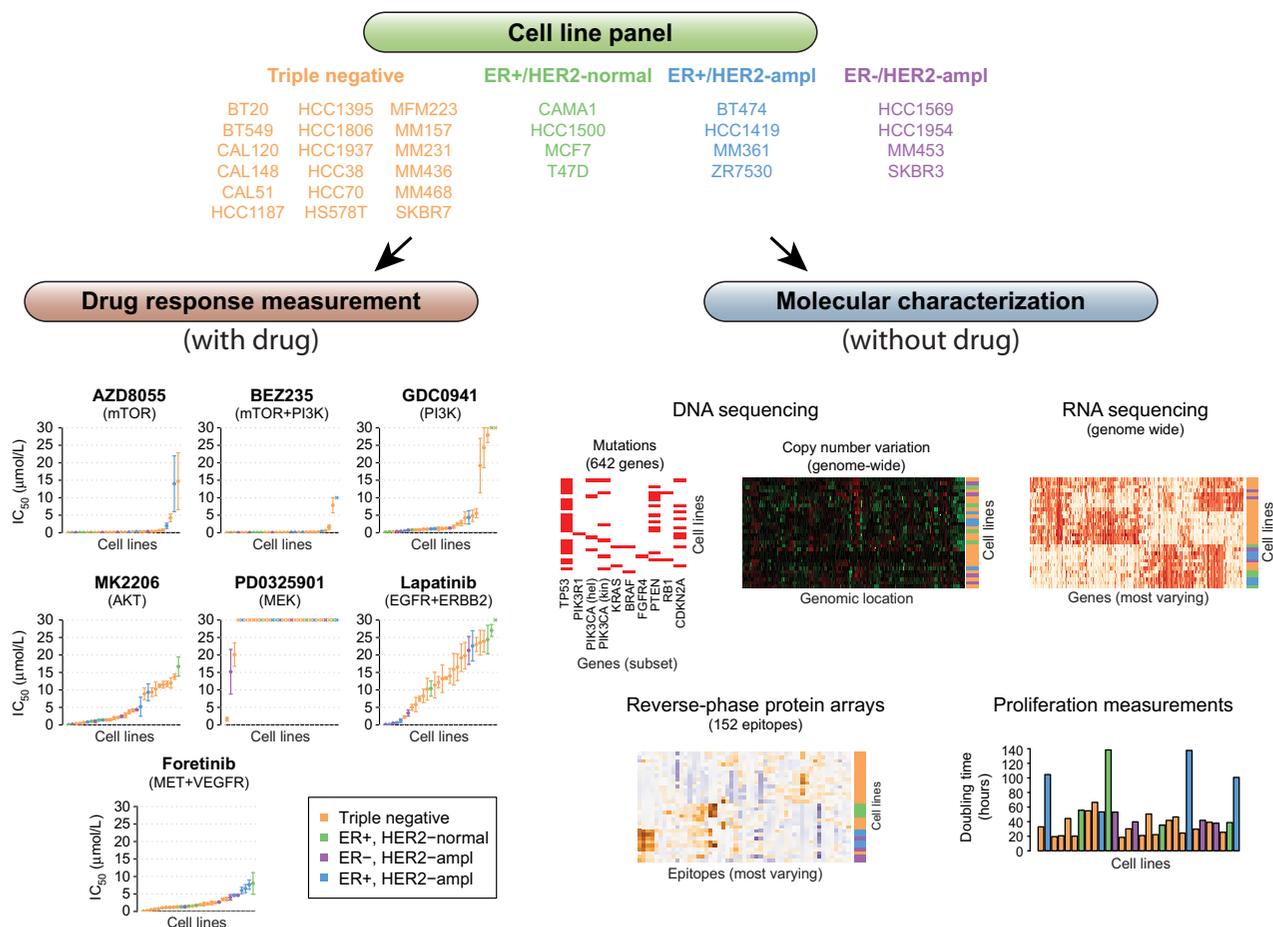
An overview of the computational model and modeling procedure is given in the Quick Guide to Equations and Assumptions, and visualized in Fig. 1. The model inference was done using BCM (19). All files required for running the inference are included in Supplementary File S1. A detailed description of all equations, data preprocessing, and inference algorithms is given in the Supplementary Materials and Methods.

Results

Establishing and characterizing a breast cancer cell line panel

We set out to establish an integrative computational model capable of explaining observed therapeutic responses based on molecular measurements. To this end, we sourced and comprehensively characterized 30 breast cancer cell lines (Fig. 2; Supplementary Table S1). Given the need for targeted treatment options for the triple-negative breast cancer subtype, the panel was enriched for triple-negative cell lines (18), with four ER⁺, four HER2⁺, and four ER⁺/HER2⁺ cell lines included to represent the other major subtypes.

The panel was characterized for response to seven kinase inhibitors, including AZD8055 (mTOR inhibitor), BEZ235 (dual mTOR/PI3K inhibitor), GDC0941 (PI3K inhibitor), MK2206 (AKT inhibitor), PD0325901 (MEK inhibitor), lapatinib (dual EGFR/HER2 inhibitor), and foretinib (cMET/VEGFR2 inhibitor). The sensitivity of each cell line to these inhibitors was determined

**Figure 2.**

Schematic of the composition and characterization of the panel of breast cancer cell lines. Thirty breast cancer cell lines were sourced and expanded, representing four major classes of breast cancer subtypes—eighteen triple-negative, four ER⁺, four HER2⁺, and four HER2⁺/ER⁺ cell lines. These cell lines were then assayed for their response to seven kinase inhibitors (bottom left, summary of response data with respect to the IC₅₀ metrics per inhibitor and cell line) as well as characterized on a molecular level using DNA capture and mutation sequencing, RNA sequencing, proteomics (RPPA analysis), and growth rate assays. In the figures and all supplementary data, abbreviated cell line names are used, in particular, MM stands for MDA-MB. Expanded views of the data plots are included in Supplementary Figs. S1 and S4–S8.

in 10-point, 72-hour dose–response assays, in biological triplicate and technical quadruplicate (summarized by IC₅₀ values in Fig. 2; Supplementary Fig. S2). The drug sensitivity measurements largely agree with those obtained in the Genomics of Drug Sensitivity in Cancer screen (Supplementary Fig. S3; refs. 10, 17, 20). However, our use of a more focused panel of 30 cell lines and seven drugs allowed us to obtain more precise measurements (Supplementary Fig. S4).

In addition to response data, we profiled the panel for mutation and copy number by DNA-seq, RNA expression by RNA-seq, protein expression, and phosphorylation by RPPA as well as proliferation rate under untreated, steady-state growth conditions (Fig. 2, bottom right; see Supplementary Figs. S5–S9, for enlarged versions of the graphs). This molecular characterization was done in the absence of drug treatment. The cell lines harbor a range of genetic events that occur in breast tumors and are present at comparable frequencies (see Supplementary Table S13, for a comparison of mutation frequencies with tumors from The Cancer Genome Atlas; ref. 2). This

cell line panel thus represents a relevant model system for the genetic diversity in breast tumors.

Fitted model provides estimates of regulatory signaling based on all available data types

To first understand which signaling is relevant for each drug, we developed a modeling framework, inference of signaling activity (ISA), to infer the signal strengths and signaling activities from all available data, as described in the Quick Guide to Equations and Assumptions and further detailed in the Supplementary Materials and Methods. We constructed a literature-based model and first fitted this to the response data for each drug separately, in conjunction with the molecular data measured in untreated cells. Although all of the interactions included in the model are well documented (Supplementary Tables S9 and S10), their relative contribution or significance is not known. For example, activating mutations in PIK3CA, the loss of PTEN, or the expression of growth factors can all lead to activation of the PI3K pathway. However, it is unclear whether

their effects are equally important, and if not, which of them has a stronger effect in a particular context.

Figure 3A illustrates the model estimates of signaling strengths (the links between signaling molecules) for lapatinib treatment. Values of the strength parameters indicate which signaling connections are important for propagating an oncogenic signal down to the proliferation node. For example, the link between ERBB2 amplification and ERBB2 activation has a strong peak at nonzero values (the density plot of $ERBB2_{amp} \rightarrow ERBB2$), thus indicating that the ERBB2 amplification gives rise to a proliferative signal. It is well known that amplification of ERBB2 and the resulting overexpression and auto-activation of this receptor provides a strong proliferation signal (21), and that this signal can be inhibited by lapatinib (22). The model provides estimates for the downstream signaling (e.g., indicating that ERBB2 signals more to PI3K than to CRAF) and for the contribution of each of the resistance mechanisms. PIK3CA mutations indeed contribute to the proliferative signal ($PIK3CA_{helical} \rightarrow PI3K$), and the model predicts that PIK3R1 mutations may have a similar effect, as we see that the parameter describing how much a PIK3R1 mutation activates PI3K tends toward higher values, and is most likely nonzero ($PIK3R1 \rightarrow PI3K$). However, the uncertainty in this parameter is large, because there is only one cell line that carries such a mutation, and consequently this parameter is only weakly constrained. As a last example, the contribution of HGF autocrine signaling (9) is represented by the parameter controlling how strongly expression of HGF leads to activation of the MET receptor ($HGF \rightarrow MET$). The posterior probability distribution of this parameter closely follows the prior, indicating that this parameter, and thus the importance of this potential resistance mechanism, cannot be determined from the current data. Together, this shows that the ISA modeling approach can be used to infer the contribution of different components driving sensitivity and resistance from all available data, while taking into account whether the parameters are identifiable.

We can use the estimates of the signaling activities (values of the nodes) to further explore the difference in signaling flow and drug response between cell lines. Figure 3B shows the estimates for ERBB2 activity and PIP3 activity in the lapatinib-treated condition scattered against the untreated condition. From the left panel, it can be seen that only the eight ERBB2-amplified cell lines show ERBB2 signaling activity, and that this activity is reduced upon lapatinib treatment in all these cell lines. In the right panel, we can see that in the lapatinib-sensitive cell lines, especially the most sensitive ones BT-474, SK-BR-3, and ZR-75-30, the reduction in ERBB2 activity also leads to a strong reduction in the PIP3 signal, whereas in the other cell lines, PIP3 signal persists, especially for the lapatinib-resistant cell line HCC1569 (see Supplementary Fig. S2, for the drug sensitivity estimates). Two non-ERBB2-amplified cell lines also have a reduced PIP3 signal upon lapatinib treatment, including T-47-D and HCC1806, which stems from their inferred EGFR activity. This illustrates the utility of the model, given that there were no molecular measurements collected in the treated conditions, and these signaling estimates are inferred from the untreated molecular data combined with the relative viability data in the treated condition. A comparison of the inferred model estimates with molecular measurements in treated condition is described later, after all model adaptations have been considered.

As a second example, for the mTOR inhibitor AZD8055, we find several factors that are associated with response. As expected, PIK3CA mutations are strongly activating in this

context and cell lines are apparently dependent on this activation (Fig. 3C, $PIK3CA_{kinase/helical} \rightarrow PI3K$ and $mTORC1 \rightarrow proliferation$), which has previously been shown (23). In addition, we find that MYC activation, as a result of gene amplification, can provide a resistance mechanism to this mTOR inhibitor ($MYC_{amp} \rightarrow proliferation$), providing another validation of our approach (24).

To facilitate the further exploration of all the signaling estimates, we generated an interface that is available at http://ccb.nki.nl/software/BCCL_KI_response_model/. This website displays the signaling strengths in each cell line upon exposure to drugs. Figure 3D shows an example of BT-474 treated with lapatinib. In this case, the model indicates that, for example, the MAPK pathway is barely involved, that there is no drug resistance provided by this cell line's RPS6KB1 amplification, and that instead, lapatinib mainly inhibits the PI3K pathway.

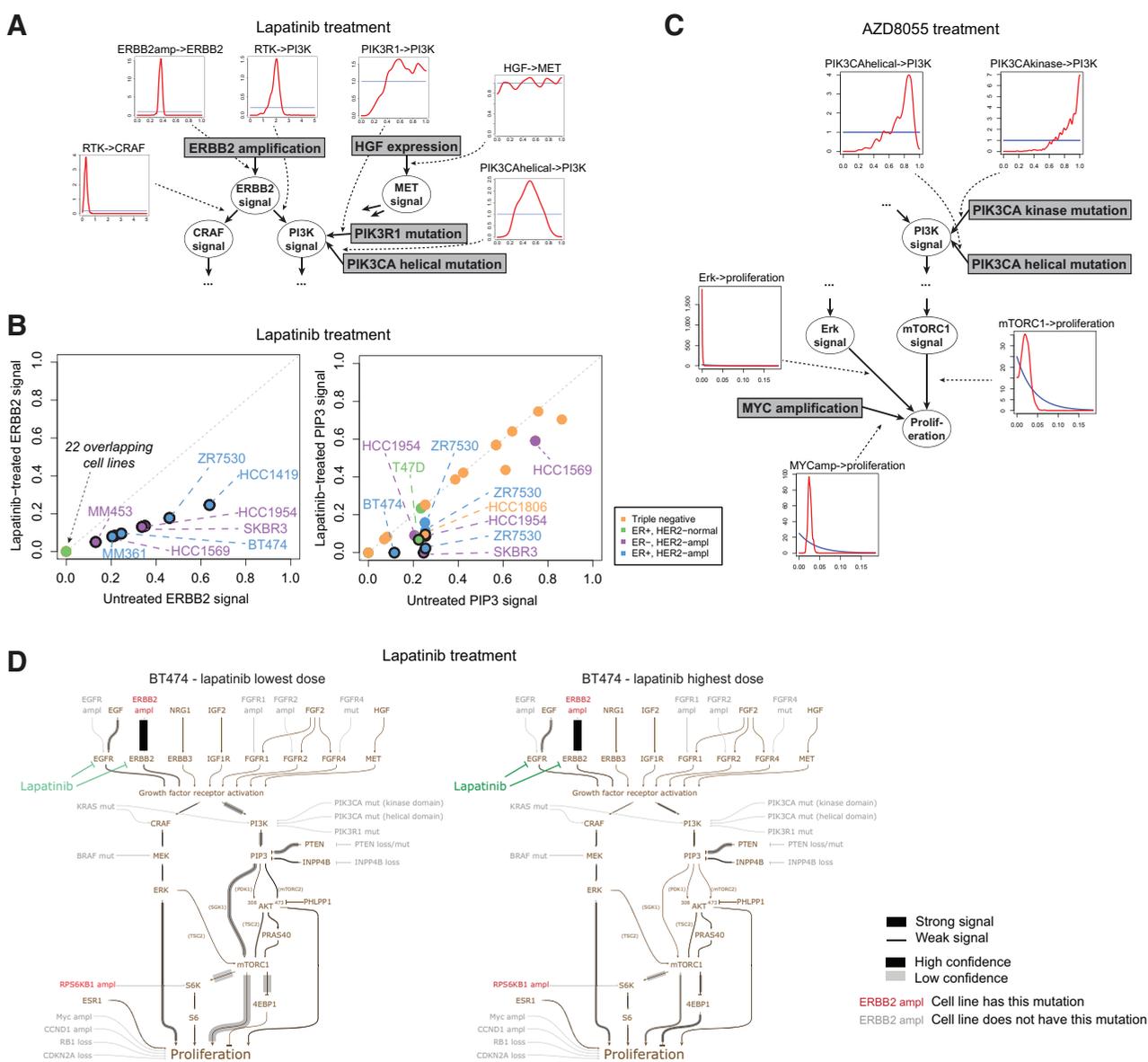
Using the posterior predictive distribution to test the goodness of fit

Although the signaling estimates appear to be reasonable, it is useful to have a systematic test of how well the fitted model describes the data. For this we used the posterior predictive distribution, which describes a new, predicted dataset based on the fitted model. We can overlay this predicted dataset on the observed measurements to have a convenient way of identifying which measurements can and cannot be explained by the model. Note that the posterior predictive distribution is not used as a measure of out-of-sample prediction; that is, it does not test how well the model predicts the behavior of unseen cell lines. Instead, it is used as a measure of goodness of fit, allowing an exploration of whether the model can describe the behavior of the cell lines at hand.

Fig. 4A shows the posterior predictive distribution overlaid on the measurement data for lapatinib. It is clear that the present model can accurately describe the relative proliferation of the cell lines as a function of drug concentration for almost all cell lines. For example, the sensitivity of the ERBB2-amplified line BT-474 and the resistance of the PIK3CA-mutated and ERBB2-amplified line MDA-MB-361 (MM361) can both be recapitulated by the model. Overviews of all posterior predictive checking for the drug response and phosphorylation data are supplied in Supplementary Data S1.

Searching for additional explanatory factors of drug sensitivity reveals novel associations

Although the model explained most of the drug response variability for lapatinib, we noticed that for some drug-cell line combinations, the fit was not as precise. For example, for foretinib, an inhibitor of c-Met and VEGFR2, we noticed that one cell line in particular, MFM-223, was much more sensitive than the model could describe (Fig. 4B). We therefore investigated the experimental data to find a possible reason for this discrepancy. A discrepancy in a single cell line is not sufficient to apply statistical tests, but we did note that this cell line has a strong FGFR2 amplification. We therefore searched the literature to see whether there is a connection and found that foretinib has in fact been reported to inhibit FGFR2 in addition to its original design targets (25). When this additional target of foretinib is added to the model, we indeed obtain a significantly improved fit (Fig. 4B). The sensitivity of other cell lines like CAL-51 and HCC-1187 to foretinib is still not explained

**Figure 3.**

Model estimates of signaling in the context of treatment with either lapatinib or AZD8055. **A**, Marginal posterior probability densities for several of the model parameters in the context of lapatinib treatment. Only the relevant parts of the model are shown. The densities are estimated using kernel density estimates with Sheather-Jones bandwidth selection. **B**, Estimates of the activity of two signaling molecules, ERBB2 and PIP3, in untreated and lapatinib-treated conditions. A black circle around a point indicates significant difference (posterior probability > 0.975 for the lapatinib-treated signal being less than the untreated signal). Error bars are not shown here; a version with error bars indicating the 90% confidence intervals is included as Supplementary Fig. S19. **C**, Like **A**, but now in the context of AZD8055 treatment. **D**, Overview of all signaling activity estimates in the BT-474 cell line, in lapatinib-treated conditions with low and high dose. Thickness of the signaling bars indicates the strength of the signal, and the scale of gray indicates uncertainty.

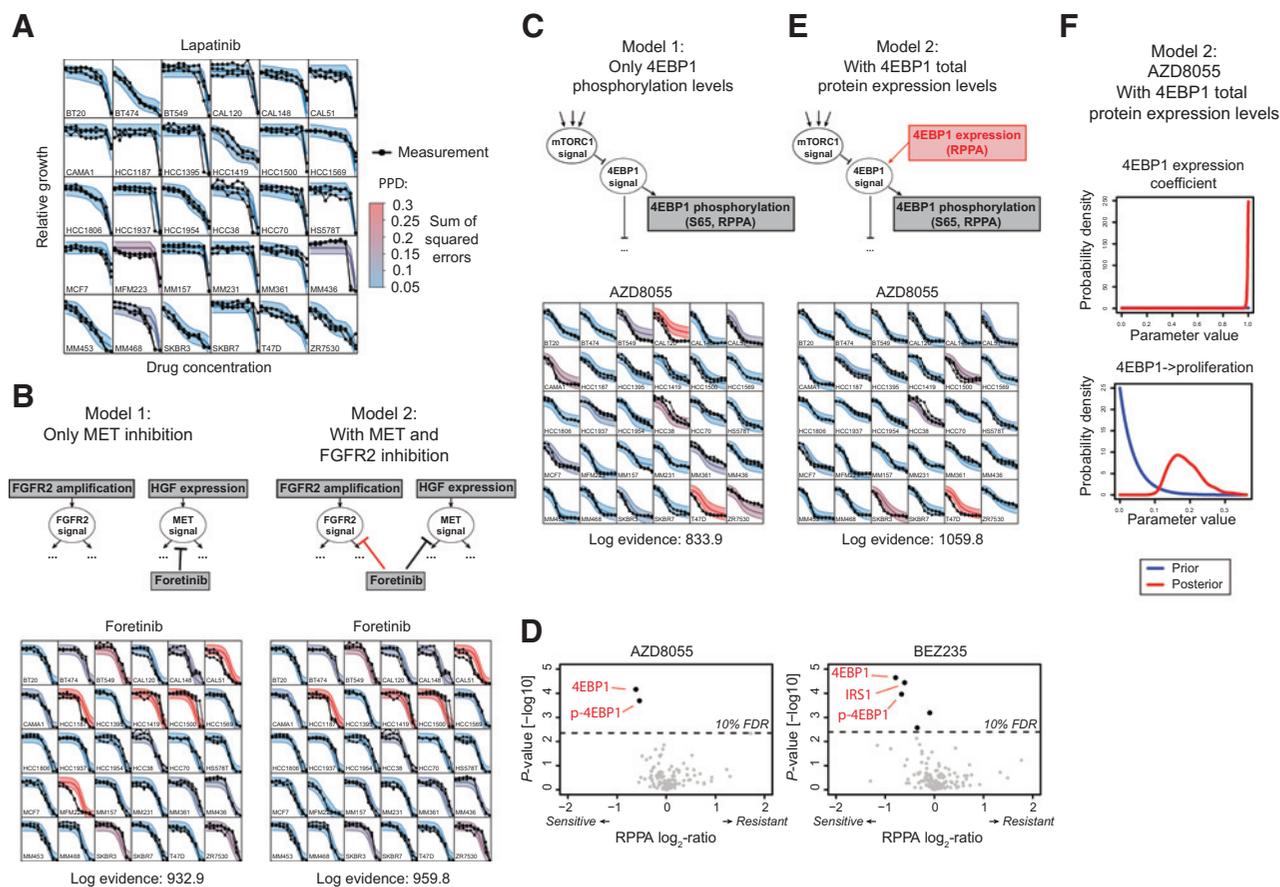
exactly, but we have not found other potential explanations for this in the data or literature, and therefore further studies would be needed to address this.

We also noticed that the model was not able to explain the response of some cell lines to mTOR inhibitors (Fig. 4C), but were unable to find additional mechanisms in the literature that could explain these discrepancies. Several cell lines are sensitive to these inhibitors even though they do not possess any of the factors known to cause sensitivity, and conversely some cell lines are

resistant despite having such sensitizing factors. For example, although BT-549, HCC1395, and HCC1937 have all lost PTEN expression, only BT-549 is sensitive to BEZ235 treatment.

We therefore further interrogated the dataset to find additional drug sensitivity mechanisms with which we could extend the model to better explain the experimental observations. With multiple sensitive cell lines we can use statistical tests. We divided the cell line panel in groups of sensitive and resistant lines using Gaussian mixture modeling, and tested whether any genes were

Jastrzebski et al.

**Figure 4.**

Goodness-of-fit testing and model expansions. **A**, The 90% confidence interval of the posterior predictive distribution for lapatinib drug response (shaded area) is overlaid on the measurement data (black). The posterior predictive distribution for each cell line is colored by the discrepancy with the data, quantified using the sum of squared errors over the eight lowest concentrations (the two highest concentrations were excluded because discrepancies at such high concentrations are likely due to various off-target effects). **B**, Comparison of two iterations of the model when fitting the foretinib drug response. **C**, Iterations of the model without protein expression levels as explanatory factor when fitting the AZD8055 drug response. **D**, Volcano plot for the association of the RPPA data with sensitivity estimates to the two mTOR inhibitors. For both inhibitors, 4E-BP1 protein expression levels correlate significantly with sensitivity. **E**, New model iteration with protein expression is included as explanatory factor. **F**, Posterior probability of two parameters controlling 4E-BP1 signaling in the expanded model. The 4E-BP1 expression coefficient (top) controls how strongly the 4E-BP1 protein expression level limits the signal through 4E-BP1. The 4E-BP1→proliferation parameter (bottom) controls how strongly the 4E-BP1 signal affects proliferation.

differentially expressed between these groups at either the RNA or protein level. The full lists of differentially expressed genes for all drugs are given in Supplementary Table S14. To further filter the differentially expressed genes, we also calculated their distance to the signaling molecules included in the model using protein-protein interaction networks (26). This provided potential candidate regulators that are not only differentially expressed, but are also functionally closely related to the signaling molecules in the model (listed in Supplementary Table S14). For both mTOR inhibitors (AZD8055 and BEZ235), the protein expression level of 4E-BP1, in addition to 4E-BP1 phosphorylation level, showed the strongest differential expression (Fig. 4D). At the RNA level, differential expression of *EIF4EBP1* was also associated with BEZ235 response. Together, these data indicated that cell lines with high expression of this protein are more sensitive to mTOR inhibitors (Supplementary Fig. S10).

To test whether the inclusion of this factor provides a better explanation for the sensitivity of some of the lines to mTOR

inhibitors, we expanded the model to include the protein expression levels of 4E-BP1. Although 4E-BP1 as a node was already included as a downstream target of mTORC1 in our previous models, only its phosphorylation state, not its protein expression level, was taken into account. Specifically, in Eq. A, the variable $E_{i,j}$ was previously based only on the binarized RNAseq expression data (Supplementary Fig. S11), and we modified this to include the RPPA protein expression levels (Supplementary Materials and Methods for details). Fig. 4E shows the model with the protein expression levels of 4E-BP1 included, whereas Fig. 4C shows the results without 4E-BP1 included. The posterior predictive checking (bottom panel) clearly shows that the expanded model provides an improved fit to the data for multiple cell lines (especially for CAL-120, BT-549, CAMA-1, and ZR-75-30), while not compromising the fit of other cell lines. The log Bayes factor between the two models (the difference in log evidence) is 226 in favor of the expanded model, where a log Bayes factor greater than 5 indicates very strong evidence (27). This indicates that the

difference is highly significant and that the improvement in fit is not merely the result of adding more free parameters. Fig. 4F shows the posterior probability of two parameters of 4E-BP1 signaling in the expanded model. The 4E-BP1 expression coefficient, shown in the top panel, describes how strongly the protein expression level of 4E-BP1 affects the amount of signal transmitted. Because the value of this parameter is very high, it indicates that the total protein expression level is a strong limiting factor for 4E-BP1 signaling in response to mTOR inhibitors. The bottom panel shows the parameter controlling the strength of 4E-BP1 signaling to proliferation, and as this is also found to be nonzero with high certainty, it implies that the 4E-BP1 signal is important for determining proliferation rate under mTOR inhibitor treatment. The computational analysis therefore predicts that the protein expression level of 4E-BP1 is an important factor in explaining mTOR inhibitor sensitivity, in addition to the already known factors determining sensitivity and resistance.

The model uses pretreatment molecular data and measurements of relative viability after drug treatment to infer signaling activities after drug treatment. To gain more confidence in the signaling estimates, we compared the estimates with measurements of protein phosphorylation of cells while on treatment (28), and found that they generally agree (Supplementary Note S1; Supplementary Fig. S12). We also constructed a reduced model, which could be fitted to the response data of all seven inhibitors at the same time (Supplementary Note S2; Supplementary Fig. S13). Finally, using an extended version of the modeling formalism (29), we confirmed that even though feedback signaling is likely to be active, the inclusion of such feedback loops does not affect how well the variability in drug response can be described (Supplementary Note S3; Supplementary Fig. S14).

The protein expression level of 4E-BP1 is a determinant of mTOR inhibitor sensitivity

Intrigued by the model prediction that 4E-BP1 protein expression is associated with mTOR inhibitor response, we investigated the biological effect of 4E-BP1 expression directly. For this, we turned to a subset of our panel of breast cancer cell lines that showed differences in response to AZD8055 and BEZ235 (Fig. 5A). These included three of the most mTOR inhibitor sensitive cell lines (BT-549, CAL-120 and Hs 578T) all bearing a gain in the *EIF4EBP1* gene-containing genomic region (Supplementary Fig. S15), which also express high levels of 4E-BP1 protein (Fig. 5B; Supplementary Fig. S16), and two insensitive cell lines (HCC1806 and HCC1937) that do not harbor a gain of the *EIF4EBP1* locus and express low levels of 4E-BP1 protein. Given that high 4E-BP1 expression may drive cells to recalibrate signaling in the pathway by increasing the expression and/or activity of mTOR, we investigated this possibility further. We first checked whether expression of 4E-BP1 and mTOR were correlated (Supplementary Fig. S17A). Although three of the most highly 4E-BP1-expressing cell lines do show an increase in mTOR expression at the protein level, in the lines chosen for our functional studies, only CAL-120 shows elevated mTOR expression. The remaining four lines (BT549, HS578T, HCC1806, HCC1937) show comparable expression of mTOR, despite vast differences in both 4E-BP1 expression and response to mTOR inhibitors. At the RNA level (Supplementary Fig. S17B), these associations were lost, suggesting that a concurrent posttranscriptional upregulation of mTOR is not a general mechanism of mTOR inhibitor sensitivity in 4E-BP1-overexpressing cells. We

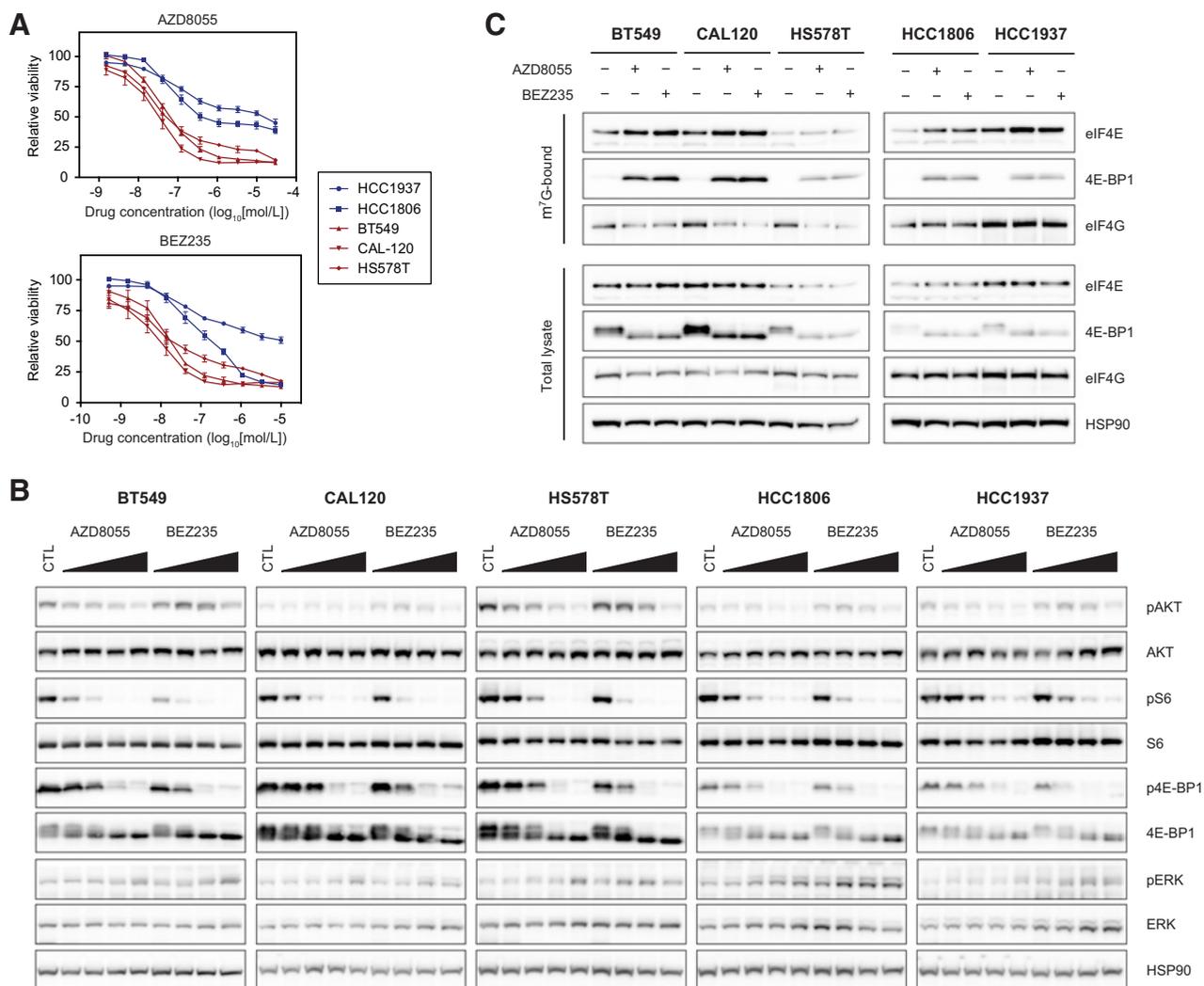
then investigated mTOR signaling in more detail by analyzing the phosphorylation and protein expression levels of several members of the PI3K, mTOR, and MAPK pathways following 24-hour treatment with two mTOR inhibitors, AZD8055 and BEZ235, in these five cell lines. This showed that both compounds had effective on-target activity, leading to reduced phosphorylation of AKT (S473), S6 (S235/236), and 4E-BP1 (S65) across all five lines, at similar compound concentrations (Fig. 5B). A minor compensatory increase in ERK phosphorylation was detected following inhibitor treatment. These data suggest that the difference in mTOR inhibitor sensitivity between these five cell lines is not caused by a difference in the compounds' ability to inhibit mTOR signaling in these lines.

To uncover the mechanism via which 4E-BP1 protein expression levels could affect response to mTOR inhibitors, we investigated the effect of inhibitor treatment on the formation of the eIF4F translation initiation complex (extensively reviewed in refs. 30, 31). As illustrated in the schematic in Supplementary Fig. S18A, the eIF4F translation complex is composed of the eIF4E and eIF4G proteins, among others. 4E-BP1 is known to negatively regulate this complex by binding and sequestering the eIF4E subunit. This displaces eIF4G from binding to eIF4E and as a result the eIF4F complex cannot initiate cap-dependent translation. The sequestering of eIF4E by 4E-BP1 is, however, inhibited when 4E-BP1 is phosphorylated by mTORC1 on several sites, which is the case when nutrients and growth factors are not limiting. Under nutrient or growth factor depletion, or alternatively following treatment with mTOR inhibitors, 4E-BP1 becomes dephosphorylated, binds to eIF4E, and thus eIF4F complex activity is repressed (Supplementary Fig. S18B). This leads to the inhibition of translation, most acutely for mRNAs with complex 5' untranslated regions (UTR) that include proliferation, survival, and tumor-promoting genes.

We investigated the dynamics of these interactions in the three mTOR inhibitor sensitive and two insensitive cell lines using the m⁷G-cap pull-down assay, which allows the visualization of the changes in eIF4G or 4E-BP1 binding to eIF4E following treatment, as compared with their expression in the total protein lysate (Fig. 5C). Our results show that mTOR inhibitor treatment leads to an increase in the binding of 4E-BP1 to eIF4E in all five cell lines, irrespective of their mTOR inhibitor response profile. In the sensitive cell lines, this increase in 4E-BP1 binding was sufficient to decrease eIF4G binding to eIF4E, as expected. In the insensitive cell lines though, the binding of 4E-BP1 to eIF4E was unable to displace eIF4G from eIF4E. This suggests that 4E-BP1 protein expression in the insensitive cell lines is below a critical threshold needed to effectively inhibit eIF4F complex formation following mTOR inhibitor treatment (Supplementary Fig. S18C), and likely explains the difference in mTOR sensitivity between these two sets of cell lines.

To further investigate whether an increase in 4E-BP1 protein expression is sufficient to increase mTOR inhibitor response, we used a lentiviral vector to overexpress 4E-BP1 in the two insensitive cell lines, HCC1806 and HCC1937 (Fig. 6A). The 4E-BP1 protein, as well as a GFP control, was effectively overexpressed in both lines and the former was detectably phosphorylated. The expression of either protein did not affect the proliferation of these cell lines, suggesting that the activity of overexpressed 4E-BP1 was efficiently inhibited by mTORC1-mediated phosphorylation (Fig. 6B). We next tested the impact of increased 4E-BP1 protein expression on the

Jastrzebski et al.

**Figure 5.**

Cell lines overexpressing the 4E-BP1 protein are more sensitive to mTOR inhibitors despite similar extent of target inhibition. **A**, Dose–response assays (72 hours) to AZD8055 (mTOR inhibitor) and BEZ235 (dual mTOR/PI3K inhibitor) in cell lines that overexpress the 4E-BP1 protein, BT-549, CAL-120, and Hs 578T (red lines), and those that do not, HCC1806 and HCC1937 (blue lines). Data represent three independent replicates \pm SEM. **B**, Western blotting results of a number of PI3K/MAPK pathway components examined across the subpanel of high- and low-4E-BP1-expressing cell lines. Twenty-four hours after seeding, cells were treated with increasing doses of AZD8055 (12, 37, 111, and 333 nmol/L) and BEZ235 (4, 12, 37, and 111 nmol/L) for a further 24 hours, whereas untreated samples served as controls (CTL). Representative of three independent experiments is shown. **C**, Cap pull-down assays to assess the effect of mTOR inhibitor treatment on the formation of the translation initiation complex, eIF4F. Twenty-four hours after seeding cells treated for a further 24 hours with 111 nmol/L AZD8055, 37 nmol/L BEZ235, or left untreated as control. Following lysis, cells were analyzed by m⁷G-cap pull-down assays to determine effects on eIF4F complex formation. Protein expression of the components in the total lysates were also determined. Representative of three independent experiments is shown.

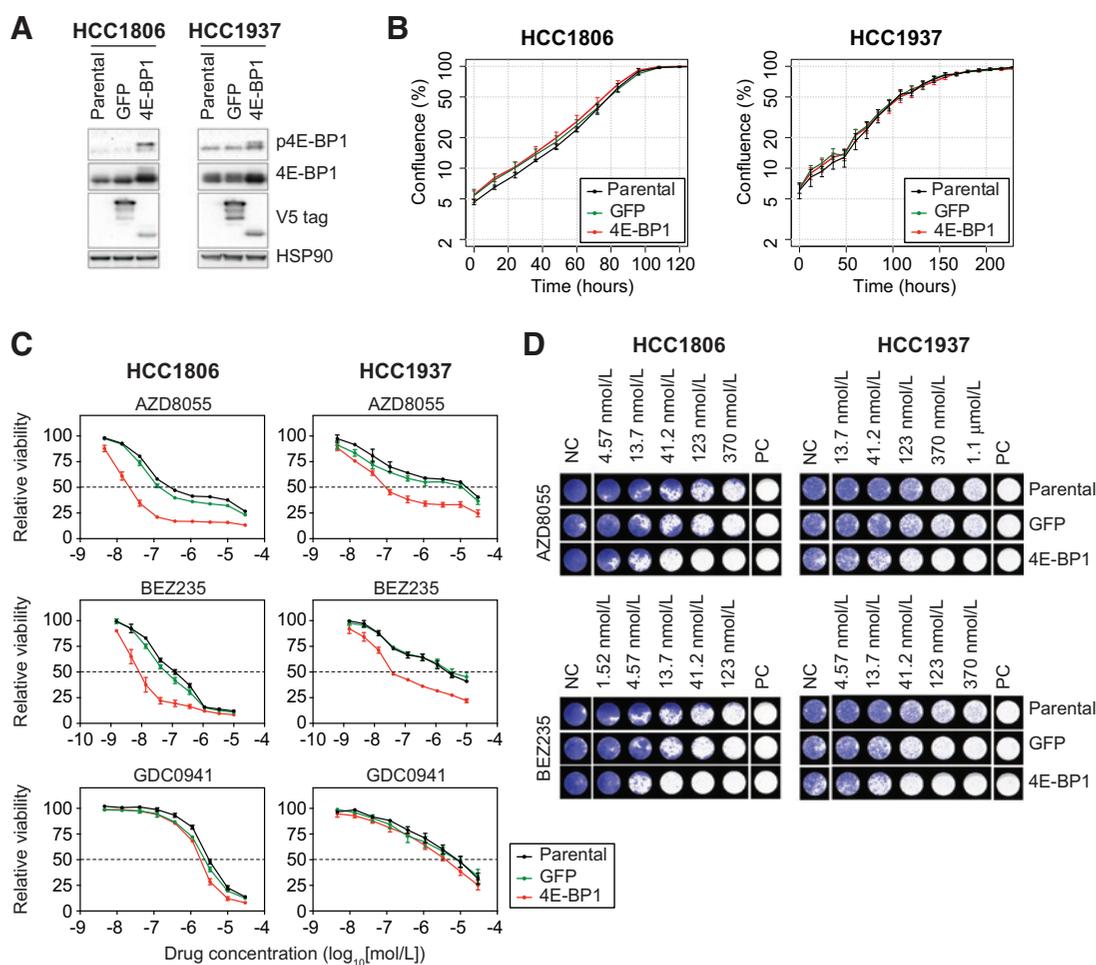
sensitivity of the cell lines to the mTOR inhibitor AZD8055, the dual mTOR/PI3K inhibitor BEZ235, as well as the PI3K inhibitor GDC0941. As shown in Fig. 6C, in short-term 72-hour drug treatment assays, the 4E-BP1-overexpressing cell lines were markedly more sensitive to mTOR inhibitors, responding at lower drug concentrations and with a decreased overall survival at higher drug concentrations. In contrast to the mTOR inhibitors, sensitivity of the 4E-BP1-overexpressing cell lines to the PI3K inhibitor GDC0941 was not increased, implying that it is specifically the inhibition of mTOR activity that is beneficial in improving response of highly expressing 4E-BP1 cell lines. We were also able to validate that 4E-BP1 overexpression increased

sensitivity to mTOR inhibitors over a longer treatment period, namely 10 days, as shown in Fig. 6D.

Together, the above results show that the level of 4E-BP1 protein expression is a determinant of sensitivity to mTOR inhibitors in breast cancer cell lines, and illustrates the utility of our computational model in identifying novel determinants of drug response in cell lines.

Discussion

Cell line panels have the potential to provide us with a better understanding of the variability in drug response between

**Figure 6.**

Overexpression of 4E-BP1 in HCC1806 and HCC1937 cell lines is sufficient to increase their sensitivity to mTOR inhibitors. **A**, Western blotting of lysates from HCC1806 and HCC1937 cell lines stably overexpressing a 4E-BP1 construct from the CCSB-Broad Lentiviral Expression Collection as compared with the parental cell lines and a GFP-overexpressing controls. **B**, Proliferation assay of the 4E-BP1-overexpressing lines as compared with the parental and GFP-expressing controls. **C**, Dose-response assays (72 hours) to AZD8055 (mTOR inhibitor), BEZ235 (dual mTOR/PI3K inhibitor), and GDC0941 (PI3K inhibitor) in the parental, GFP- and 4E-BP1-expressing HCC1806 and HCC1937 cell lines. Data represent three independent replicates \pm SEM. **D**, Long-term (10 day) dose-response assay to AZD8055 and BEZ235 in the parental, GFP- and 4E-BP1-expressing HCC1806 and HCC1937 cell lines. Representative of three independent experiments is shown.

patients. Previous efforts of linking molecular characteristics to drug sensitivity in cell line panels have identified several known and novel associations (10, 11, 13–16). Here we showed that by combining extensive measurements with mathematical modeling, a more detailed understanding of variability in drug sensitivity can be achieved. The use of Bayesian statistical analysis allowed for the simultaneous integration of diverse data types with prior knowledge from the literature.

Models are by definition a simplified representation of the system. Two major simplifications that were used here are the assumption of quasi-steady state and the absence of feedback signaling. The benefit of using these simplifications is that significantly more components of cellular signaling can be included in the model. It is reassuring that a model with these simplifications can describe a large part of the variability in short-term drug response. Studying longer-term drug response and, for example, adaptive resistance will likely require the incorporation of dynam-

ics and feedback mechanisms. Using dynamical models or the inclusion of feedback mechanisms does not pose any theoretical problems for the modeling approach we used. However, the computational cost is significantly higher, and additional intervention or time-course data would be needed to constrain the parameters.

Recently, Fey and colleagues described a computational model of JNK signaling, containing five signaling proteins that provided prognostic information in patients with neuroblastoma (32), showing that such computational models may be useful also in a clinical setting. Their model was informative in a specific subset of patients, namely those with MYCN-amplified tumors, whereas in the general population individual biomarkers were still more informative. This indicates that it is necessary for models to incorporate various different oncogenic drivers, residing in multiple signaling pathways, to capture the variability across a wide range of patients.

A method that integrates multiple data types to obtain pathway activation status is PARADIGM (33). Heiser and colleagues have used a modified version of this method, SuperPathway analysis, to link pathway activation status to drug response in a large breast cancer cell line panel (12). They found, among others, that upregulation of DNA damage response pathways was associated with sensitivity to cisplatin, although this was not further tested experimentally. Indeed, PARADIGM and SuperPathway analyses do not shed light on how this association might work mechanistically, and the involvement of individual components of the DNA damage response pathway, such as *TP53*, *ATM*, and *BRCA1/2*, is not investigated. In contrast, the approach presented here provides detailed signaling flows and estimates the relative contributions from all drivers and sensitivity mechanisms included in the model, making the *in silico* findings amenable to experimental validation.

In the course of refining our model, we found that elevated 4E-BP1 protein expression played an important role in the response of breast cancer cell lines to AZD8055 and BEZ235, two inhibitors targeting the mTOR kinase. We further validated this observation in 4E-BP1 overexpression studies, showing that this provides a pool of an endogenous translation inhibitor available for activation, and thus inhibition of cap-dependent translation, via mTOR inhibitor treatment.

On a mechanistic level, these findings are in line with prior work using transformed mouse embryonic fibroblasts, which showed that the ratio of eIF4E/4E-BP1 expression can predict response to mTOR-directed therapy (34). That is, a higher ratio of eIF4E/4E-BP1 predicted poorer response to mTOR inhibitors. Consistent with this, *EIF4E* amplification was reported as a mechanism of AZD8055 resistance in a SW620 colorectal cell line model (35). Conversely, a lack of 4E-BP1 expression in lymphoma cells, thus a reduced ability to restrain eIF4E activity, has been shown to lead to resistance to mTOR inhibition, an effect reversed by exogenous expression of the 4E-BP1 protein (36), similar to our findings in 4E-BP1-overexpressing HCC1806 and HCC1937 cell lines. An exception to these findings is the report of elevated 4E-BP1 protein levels in mTOR inhibitor treatment-resistant luminal subpopulation of prostate cancer cells (37), although it remains to be investigated whether this observation is restricted to prostate cancer or the luminal subtype. Most recently, a study by Wang and colleagues (38) added further weight to the hypothesis that elevated 4E-BP1 expression can be a marker of mTOR inhibitor sensitivity. They show that the combination of an mTOR inhibitor and an HDAC inhibitor, the latter acting to derepress Snail-mediated 4E-BP1 transcriptional inhibition, can synergize to inhibit tumor growth in mice.

Notably, in our cell line panel, the overexpression of 4E-BP1 resulted from a copy number gain in the genomic region encoding *EIF4EBP1*. Focal amplification of the 8p11-12 region, the region containing *EIF4EBP1*, is a known event in breast cancer occurring in approximately 15% of cases, and patients who harbor this event in their primary tumor have a much higher likelihood of relapse (39). Previous studies have identified various genes in this region as potential oncogenes, with most evidence so far supporting *FGFR1* (40–42) and *ZNF703* (43–45). These two genes lie close to and on either side of *EIF4EBP1*. Our cell line experiments show that overexpression of 4E-BP1 alone does not affect viability *in vitro*. Although we cannot exclude that amplification of 4E-BP1 can contribute to a transformed phenotype in tumors (such as affecting invasion,

migration, or cell viability *in vivo*), various studies indicate that the cellular function of 4E-BP1 is consistent with a role as a tumor suppressor gene, rather than as an oncogene (34–36, 46–49). It therefore seems plausible that the amplification of *EIF4EBP1* is a passenger event. This raises the possibility that by selecting for amplification of nearby oncogenes, the tumors have also introduced a specific "passenger vulnerability"—that is, a passenger aberration that introduces a vulnerability to a particular drug. If the drug sensitivity association translates from cell lines to patients, testing for *EIF4EBP1* amplifications could identify patients who may benefit most from treatment with mTOR inhibitors such as AZD8055 and BEZ235.

One difference between the cell line and patient data is that the subtypes in which the gain of *EIF4EBP1* is present vary. Although we identified this event predominantly in the triple-negative and ER+ breast cancer cell lines, in patients, the 8p11-12 amplicon is present almost exclusively in ER+ tumors. Interestingly, phase III clinical trials with the allosteric mTOR inhibitor, everolimus, have been carried out in the ER+ setting, showing a significant improvement in progression-free survival (PFS) in patients who received everolimus versus placebo in addition to the aromatase inhibitor, exemestane (50). A subsequent analysis of this data, exploring associations of PFS with common genetic aberrations, has found no improvement in PFS in patients with *FGFR1* gene amplification (generally coamplifying the *EIF4EBP1* gene; ref. 51). Although this may suggest that there is no increased clinical benefit for mTOR inhibitors in patients with 8p11-12 amplified tumors, the inhibitory activity of everolimus, a rapalog, versus active-site inhibitors, such as AZD8055 and BEZ235, differs substantially. Although everolimus is an allosteric inhibitor of predominantly mTOR complex 1 (mTORC1), active-site inhibitors are able to target both mTORC1 and mTORC2 (30). Most importantly, the active-site inhibitors have been shown to result in a much more potent inhibition of downstream mTOR signaling, specifically with respect to the inhibition of 4E-BP1 phosphorylation, whereas in the case of rapalogs this phosphorylation is restored within hours of treatment (52, 53). As such, it is likely that the absence of an association between amplification in the 8p11-12 region containing the *FGFR1* and *EIF4EBP1* genes and PFS with everolimus treatment observed by Hortobagyi and colleagues (51) stems from limited inhibition of 4E-BP1 phosphorylation. Together, these studies suggest that the benefit of rapamycin treatment in ER+ tumors results from a mechanism distinct to that of 4E-BP1 inhibition, but also emphasize the need for further study to determine the efficacy of active-site mTOR inhibitors in patients with 8p11-12 amplifications.

We conclude that the combination of mathematical modeling of signaling in response to drug treatment with a large panel of molecularly characterized cell lines demonstrates the usefulness of combining large data sets with prior knowledge to uncover key determinants of drug sensitivity. Although the work presented here applied the ISA methodology to explain drug response in a panel of cell lines, our ultimate goal is to develop this approach into a tool for predicting response in patients. Further validation on clinical samples is of course required, but a clear benefit is that the molecular characterization, needed as input for a model, can be performed directly on biopsy material. This circumvents a need for *ex vivo* culture or lengthy response profiling, making the approach amenable to clinical application. We believe that such systematic, quantitative approaches to the understanding of drug responses hold

promise as tools for achieving the goal of precision medicine in cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: K. Jastrzebski, B. Thijssen, R.L. Beijersbergen, L.F.A. Wessels

Development of methodology: K. Jastrzebski, B. Thijssen

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): K. Jastrzebski, K. de Lint, I.J. Majewski

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): K. Jastrzebski, B. Thijssen, R.J.C. Kluin, L.F.A. Wessels

Writing, review, and/or revision of the manuscript: K. Jastrzebski, B. Thijssen, K. de Lint, R.L. Beijersbergen, L.F.A. Wessels

Study supervision: R.L. Beijersbergen, L.F.A. Wessels

Acknowledgments

We are grateful to Thomas Kuilman for performing the CopywriteR analysis and the NKI Genomics Facility for performing high-throughput sequencing experiments and subsequent data processing. We are grateful to Jordi Vidal Rodriguez for help with the drug response assays and maintenance of the cell line panel. We thank the Beijersbergen, Wessels, and Bernards groups for helpful discussions. This work was funded by the Cancer Systems Biology Center grant (853.00.120) from the Netherlands Organization for Scientific Research (NWO) and the Cancer Genomics Netherlands Program supported by the Gravitation program of the Netherlands Organization for Scientific Research (NWO).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 6, 2017; revised February 26, 2018; accepted May 21, 2018; published first May 29, 2018.

References

- Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin* 2014;64:9–29.
- The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* 2012;490:61–70.
- Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012;486:346–52.
- Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE, Davidson NE, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005;353:1673–84.
- Geyer CE, Forster J, Lindquist D, Chan S, Romieu CG, Pienkowski T, et al. Lapatinib plus capecitabine for HER2-positive advanced breast cancer. *N Engl J Med* 2006;355:2733–43.
- Groenendijk FH, Bernards R. Drug resistance to targeted therapies: Deja vu all over again. *Mol Oncol* 2014;8:1067–83.
- Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hijmans EM, Beelen K, et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 2007;12:395–402.
- Nagata Y, Lan KH, Zhou X, Tan M, Esteva FJ, Sahin AA, et al. PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* 2004;6:117–27.
- Wilson TR, Fridlyand J, Yan Y, Penuel E, Burton L, Chan E, et al. Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. *Nature* 2012;487:505–9.
- Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* 2012;483:570–5.
- Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;483:603–7.
- Heiser LM, Sadanandam A, Kuo W-L, Benz SC, Goldstein TC, Ng S, et al. Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proc Natl Acad Sci U S A* 2012;109:2724–9.
- Daemen A, Griffith OL, Heiser LM, Wang NJ, Enache OM, Sanborn Z, et al. Modeling precision treatment of breast cancer. *Genome Biol* 2013;14:R110.
- Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, et al. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. *Cancer Discov* 2015;5:1210–23.
- Kirouac DC, Du JY, Lahdenranta J, Overland R, Yarar D, Paragas V, et al. Computational modeling of ERBB2-amplified breast cancer identifies combined ErbB2/3 blockade as superior to the combination of MEK and AKT inhibitors. *Sci Signal* 2013;6:ra68.
- Klinger B, Sieber A, Fritsche-Guenther R, Witzel F, Berry L, Schumacher D, et al. Network quantification of EGFR signaling unveils potential for targeted combination therapy. *Mol Syst Biol* 2013;9:673. doi: 10.1038/msb.2013.29.
- Iorio F, Knijnenburg TA, Vis DJ, Saez-Rodriguez J, McDermott U, Garnett MJ, et al. A landscape of pharmacogenomic interactions in cancer. *Cell* 2016;166:740–54.
- Sun C, Wang L, Huang S, Heynen GJ, Prahallad A, Robert C, et al. Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. *Nature* 2014;508:118–22.
- Thijssen B, Dijkstra TMH, Heskes T, Wessels LFA. BCM: toolkit for Bayesian analysis of computational models using samplers. *BMC Syst Biol* 2016;10:100.
- Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* 2013;41:D955–61.
- Arteaga CL, Engelman JA. ERBB receptors: From oncogene discovery to basic science to mechanism-based cancer therapeutics. *Cancer Cell* 2014;25:282–303.
- Rusnak DW, Lackey K, Affleck K, Wood ER, Alligood KJ, Rhodes N, et al. The effects of the novel, reversible epidermal growth factor receptor/ErBB-2 tyrosine kinase inhibitor, GW2016, on the growth of human normal and tumor-derived cell lines in vitro and in vivo. *Mol Cancer Ther* 2001;1:85–94.
- Serra V, Markman B, Scaltriti M, Eichhorn PJ, Valero V, Guzman M, et al. NVP-BE2255, a dual PI3K/mTOR inhibitor, prevents PI3K signaling and inhibits the growth of cancer cells with activating PI3K mutations. *Cancer Res* 2008;68:8022–30.
- Ilic N, Utermark T, Widlund HR, Roberts TM. PI3K-targeted therapy can be evaded by gene amplification along the MYC-eukaryotic translation initiation factor 4E (eIF4E) axis. *Proc Natl Acad Sci* 2011;108:E699–708.
- Kataoka Y, Mukohara T, Tomioka H, Funakoshi Y, Kiyota N, Fujiwara Y, et al. Foretinib (GSK1363089), a multi-kinase inhibitor of MET and VEGFRs, inhibits growth of gastric cancer cell lines by blocking inter-receptor tyrosine kinase networks. *Invest New Drugs* 2012;30:1352–60.
- Das J, Yu H. HINT: High-quality protein interactomes and their applications in understanding human disease. *BMC Syst Biol* 2012;6:1–12.
- Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc* 1995;90:773–95.
- Korkola JE, Collisson EA, Heiser L, Oates C, Bayani N, Itani S, et al. Decoupling of the PI3K pathway via mutation necessitates combinatorial treatment in HER2+ breast cancer. *PLoS One* 2015;10:e0133219.
- Thijssen B, Jastrzebski K, Beijersbergen RL, Wessels LFA. Delineating feedback activity in the MAPK and AKT pathways using feedback-enabled inference of signaling activity. *bioRxiv* 2018. doi: <https://doi.org/10.1101/268359>.
- Bhat M, Robichaud N, Hulea L, Sonenberg N, Pelletier J, Topisirovic I. Targeting the translation machinery in cancer. *Nat Rev Drug Discov* 2015;14:261–78.
- Pelletier J, Graff J, Ruggero D, Sonenberg N. Targeting the eIF4F translation initiation complex: a critical nexus for cancer development. *Cancer Res* 2015;75:250–63.

Jastrzebski et al.

32. Fey D, Halasz M, Dreidax D, Kennedy SP, Hastings JF, Rauch N, et al. Signaling pathway models as biomarkers: Patient-specific simulations of JNK activity predict the survival of neuroblastoma patients. *Sci Signal* 2015;8:RA130.
33. Vaske CJ, Benz SC, Sanborn JZ, Earl D, Szeto C, Zhu J, et al. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* 2010;26:237–45.
34. Alain T, Morita M, Fonseca BD, Yanagiya A, Siddiqui N, Bhat M, et al. eIF4E/4E-BP ratio predicts the efficacy of mTOR targeted therapies. *Cancer Res* 2012;72:6468–76.
35. Cope CL, Gilley R, Balmanno K, Sale MJ, Howarth KD, Hampson M, et al. Adaptation to mTOR kinase inhibitors by amplification of eIF4E to maintain cap-dependent translation. *J Cell Sci* 2014;127:788–800.
36. Mallya S, Fitch BA, Lee JS, So L, Janes MR, Fruman DA. Resistance to mTOR kinase inhibitors in lymphoma cells lacking 4EBP1. *PLoS One* 2014;9:1–10.
37. Hsieh AC, Nguyen HG, Wen L, Edlind MP, Carroll PR, Kim W, et al. Cell type-specific abundance of 4EBP1 primes prostate cancer sensitivity or resistance to PI3K pathway inhibitors. *Sci Signal* 2015;8:ra116–ra116.
38. Wang J, Ye Q, Cao Y, Guo Y, Huang X, Mi W, et al. Snail determines the therapeutic response to mTOR kinase inhibitors by transcriptional repression of 4E-BP1. *Nat Commun* 2017;8:2207. doi: 10.1038/s41467-017-02243-3.
39. Gelsi-Boyer V, Orsetti B, Cervera N, Finetti P, Sircoulomb F, Rougé C, et al. Comprehensive profiling of 8p11-12 amplification in breast cancer. *Mol Cancer Res* 2005;3:655–67.
40. Theillet C, Adelaide J, Louason G, Bonnet-Dorion F, Jacquemier J, Adnane J, et al. FGFR1 and PLAT genes and DNA amplification at 8p12 in breast and ovarian cancers. *Gene Chromosome Canc* 1993;7:219–26.
41. Ugolini F, Adelaide J, Charafe-Jauffret E, Nguyen C, Jacquemier J, Jordan B, et al. Differential expression assay of chromosome arm 8p genes identifies Frizzled-related (FRP1/FRZB) and Fibroblast Growth Factor Receptor 1 (FGFR1) as candidate breast cancer genes. *Oncogene* 1999;18:1903–10.
42. Turner N, Pearson A, Sharpe R, Lambros M, Geyer F, Lopez-Garcia Ma., et al. FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer Res* 2010;70:2085–94.
43. Garcia MJ, Pole JCM, Chin S-F, Teschendorff A, Naderi A, Ozdag H, et al. A 1 Mb minimal amplicon at 8p11-12 in breast cancer identifies new candidate oncogenes. *Oncogene* 2005;24:5235–45.
44. Holland DG, Burleigh A, Git A, Goldgraben Ma, Perez-Mancera Pa, Chin S-F, et al. ZNF703 is a common Luminal B breast cancer oncogene that differentially regulates luminal and basal progenitors in human mammary epithelium. *EMBO Mol Med* 2011;3:167–80.
45. Slorach EM, Chou J, Werb Z. Zeppo1 is a novel metastasis promoter that represses E-cadherin expression and regulates p120-catenin isoform expression and localization. *Genes Dev* 2011;25:471–84.
46. Ducker GS, Atreya CE, Simko JP, Hom YK, Matli MR, Benes CH, et al. Incomplete inhibition of phosphorylation of 4E-BP1 as a mechanism of primary resistance to ATP-competitive mTOR inhibitors. *Oncogene* 2014;33:1590–600.
47. Martineau Y, Azar R, Müller D, Lasfargues C, El Khawand S, Anesia R, et al. Pancreatic tumours escape from translational control through 4E-BP1 loss. *Oncogene* 2014;33:1367–74.
48. Cai W, Ye Q, She Q-B. Loss of 4E-BP1 function induces EMT and promotes cancer cell migration and invasion via cap-dependent translational activation of snail. *Oncotarget* 2014;5:6015–27.
49. Martín ME, Pérez MI, Redondo C, Alvarez MI, Salinas M, Fando JL. 4E binding protein 1 expression is inversely correlated to the progression of gastrointestinal cancers. *Int J Biochem Cell Biol* 2000;32:633–42.
50. Yardley DA, Noguchi S, Pritchard KI, Burris HA, Baselga J, Gnant M, et al. Everolimus plus exemestane in postmenopausal patients with HR+ breast cancer: BOLERO-2 final progression-free survival analysis. *Adv Ther* 2013;30:870–84.
51. Hortobagyi GN, Chen D, Piccart M, Rugo HS, Burris HA, Pritchard KI, et al. Correlative analysis of genetic alterations and everolimus benefit in hormone receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer: results from BOLERO-2. *J Clin Oncol* 2015;34:419–26.
52. Choo AY, Yoon S-O, Kim SG, Roux PP, Blenis J. Rapamycin differentially inhibits S6Ks and 4E-BP1 to mediate cell-type-specific repression of mRNA translation. *Proc Natl Acad Sci U S A* 2008;105:17414–9.
53. Choo AY, Blenis J. Not all substrates are treated equally: implications for mTOR, rapamycin-resistance and cancer therapy. *Cell Cycle* 2009;8:567–72.

Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

Integrative Modeling Identifies Key Determinants of Inhibitor Sensitivity in Breast Cancer Cell Lines

Katarzyna Jastrzebski, Bram Thijssen, Roelof J.C. Kluin, et al.

Cancer Res 2018;78:4396-4410. Published OnlineFirst May 29, 2018.

Updated version Access the most recent version of this article at:
doi:[10.1158/0008-5472.CAN-17-2698](https://doi.org/10.1158/0008-5472.CAN-17-2698)

Supplementary Material Access the most recent supplemental material at:
<http://cancerres.aacrjournals.org/content/suppl/2018/05/26/0008-5472.CAN-17-2698.DC1>

Visual Overview **A diagrammatic summary of the major findings and biological implications:**
<http://cancerres.aacrjournals.org/content/78/15/4396/F1.large.jpg>

Cited articles This article cites 52 articles, 16 of which you can access for free at:
<http://cancerres.aacrjournals.org/content/78/15/4396.full#ref-list-1>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link
<http://cancerres.aacrjournals.org/content/78/15/4396>.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.