Integrating network reconstruction with mechanistic modelling to predict potential cancer therapy

Melinda Halasz[1], Boris N. Kholodenko[1,2,3], Walter Kolch[1,2,3*], Tapesh Santra[1*]

[1]Systems Biology Ireland, University College Dublin, Belfield, Dublin-4, Ireland

[2]Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin, Ireland

[3]School of Medicine and Medical Science, University College Dublin, Belfield, Dublin, Ireland

[*] Corresponding authors: walter.kolch@ucd.ie, tapesh.santra@ucd.ie

Running title: A framework for pathway inference and modelling

Character count: 48856

Abstract

Signal transduction networks (STNs) are often rewired in cancerous cells. Effective cancer treatment requires identifying and repairing these harmful alterations. We developed a computational framework which can identify these aberrations and predict potential targets for intervention. It reconstructs network models of STNs from noisy and incomplete perturbation response data, and then uses the reconstructed networks to develop mechanistic models of STNs for predicting potential treatments. As a proof of principle, we analysed a perturbation dataset targeting Epidermal Growth Factor Receptor (EGFR) and Insulin like

Growth Factor 1 Receptor (IGF1R) pathways in a panel of colorectal cancer (CRC) cells, revealing cell line specific STN rewiring. Specifically, we found that the feedback inhibition of IRS1 by p70S6K is associated with resistance to EGF receptor (EGFR) inhibition, and disrupting this feedback may restore sensitivity to EGFR inhibitors in CRC cells. These findings were experimentally validated *in vitro* and in zebrafish (*Danio rerio*) xenografts.

Keywords: Network inference, mechanistic modelling, modular response analysis, signal transduction pathways, cancer, drug resistance, Bayesian inference.

Introduction

Signal transduction networks (STNs) are frequently rewired in diseases, and reconstructing their topologies and dynamic changes is challenging.[1] Most existing methods reconstruct logical or probabilistic network models, based on correlated behaviour observed in experimental data.[2-8] Such models delineate global connections, but typically do not allow predictive analyses of STN adaptations. Detailed mathematical models[9-16] of STNs permit such analyses, but require exact prior knowledge of the STN topologies and extensive quantitative experiments for parameterizations, hampering their application for investigating STN rewiring. Modular Response Analysis (MRA) takes the middle ground between qualitative and detailed mathematical modelling. [17-20] MRA reconstructs simple linear models[17], which recover STN topologies, and within restrictions can predict their functions.[20] Unfortunately, MRA is sensitive to experimental noise[21], requires large numbers of perturbation experiments (equalling the number of network components)[17], and has limited predictive power in some respects; e.g. MRA cannot predict the effects of sequential drug applications[20], or differentiate between drugs that target the same protein(s) using different molecular mechanisms.[22]

Here, we present a computational framework which overcomes these limitations. It reconstructs network models, and subsequently, derives Ordinary Differential Equation (ODE) based mechanistic models of STNs from noisy and incomplete datasets by exploiting MRA in a Bayesian setting (BMRA). The resulting ODE models are used to predict STN adaptations in response to therapeutic interventions. Testing our method on a publicly available dataset[20], revealed rewiring of the EGF and IGF-1 STNs in different colorectal cancer (CRC) cells and uncovered a previously unknown functional association between a negative feedback loop and EGFR inhibitor resistance in some of these cell lines. It further correctly predicted the responses of the EGFR and IGF1R STNs to different individual and combinatorial perturbations. Importantly, it predicted that inhibiting p70S6K helps overcoming EGFR inhibitor resistance in some CRC cells. These predictions were validated *in vitro* and in *in vivo* xenograft models.

Results

A computational framework that integrates network reconstruction methods with mechanistic modelling of STNs

Our computational framework consists of two stages which involve a network reconstruction and a mechanistic modelling algorithm, respectively (Fig. 1A). Unlike other approaches[10, 13-16, 23, 24], the mechanistic modelling algorithm in our framework inherently relies on the network models reconstructed by the first stage and does not require time course measurements of pathway components for parameter calibration. Thus, it can use relatively little experimental data to provide an integrated pipeline that allows both reconstructing and analysing STNs *in silico*. Below we provide a brief overview of our method.

Network reconstruction: We developed a Bayesian algorithm (BMRA) that builds on MRA [17, 18] to reconstruct STN topologies (Fig. 1B). Bayesian algorithms use *experimental data* to update *prior knowledge*, and, have two main components, the *prior distribution* and the *likelihood function*. The prior distribution evaluates how strongly a network model is supported by prior knowledge, whereas the likelihood function determines how well it explains experimental data. The prior distribution is formulated by assembling all known interactions of a STN into a prior network, and then formulating a probability distribution that assigns higher probabilities to networks that are topologically similar to the prior network (for details see Materials and Methods, abbreviated M&M hereafter). The likelihood function is calculated by exploiting MRA[17, 18] which establishes linear predictive models of networks [17, 18] that can be used to evaluate how likely a network is to reproduce experimentally observed data (M&M, Supp. Note 2). The product of the prior distribution and the likelihood function yields the non-normalized posterior distribution which evaluates how well a network model is supported by prior knowledge and experimental data combined (M&M, Supp. Note 2) [25]. However, the normalization constant of the posterior distribution is needed to estimate necessary statistical features of the network, e.g. the probability and average interaction strength (with confidence interval) of each network connection. But it is not straight forward to calculate this constant analytically. Therefore, Markov Chain Monte Carlo (MCMC) sampling is used to approximate the true posterior distributions from its non-normalized counterpart (M&M, Supp. Note 2). The MCMC algorithm yields the posterior probability of each interaction and probability distribution of the associated interaction strengths.

Before applying BMRA to experimental datasets, we benchmarked its performance using public benchmark datasets (from DREAM consortium; http://dreamchallenges.org/) and evaluated its sensitivity to noise and data incompleteness using simulated data. In the

benchmarking studies, it ranked between 2nd and 4th among 29 other algorithms on five benchmarking datasets when no prior information was used to infer networks (Supp. Note 2). When networks containing 50% false interactions were used as prior knowledge, BMRA performed nearly flawlessly (Supp. Note 2). The sensitivity evaluation study revealed that BMRA is a highly robust algorithm and outperformed three recent algorithms when tested on datasets with different levels of noise and data incompleteness (Supp. Note 2).

Mechanistic modelling: The reconstructed network models are fed into Bayesian mechanistic modelling (BMM) algorithm, which we have designed in this study, for developing predictive models of STNs (Fig. 1C). Firstly, a generic model of pathway reactions is generated using existing literature. An ODE based mathematical model is then developed for this generic reaction model. The mathematical model formulates the rate of change in the concentration of each pathway component by a parametric function of its own concentration and that of its regulators. The parameters of these rate functions are then fitted using a Bayesian algorithm. The efficiency and accuracy of parameter fitting algorithms decay rapidly as the number of parameters increases [26]. Therefore it is desirable to select as few model parameters as possible for fitting. We select only those parameters which have maximum impact on the model output. The impact of a parameter on the output of an STN is quantified by the fractional change in the STN output caused by a small change in the parameter (M&M). These quantities (also known as parameter sensitivities) depend on parameter values. Therefore, we estimated these quantities for a large number of random parameter values using a Monte Carlo approach (M&M). The parameters which were found to have the highest overall impact on the STN output were then selected for fitting (M&M), while keeping the remaining parameters at fixed values. We first assigned prior distributions to the selected parameters based on initial guesses about the range of their potential values (Supp. Note 3, M&M). These prior distributions were then updated using Bayes' rule[25] to

estimate posterior distributions. The posterior distribution is proportional to the product of the prior distributions and the likelihood function which estimates the likelihood that a set of parameter values will result in simulated interaction strengths which have the same distribution as inferred by the BMRA algorithm (M&M, Supp. Note 3). The proportionality constant of the posterior cannot be calculated analytically, and hence MCMC sampling is used to approximate the true posterior from its non-normalized counterpart. The posterior parameter distributions are used to simulate model behaviour.

Reconstructing EGFR and IGF1R networks in a panel of CRC cells

We tested the above analysis pipeline by tackling the problem of EGFR resistance in colorectal cancer (CRC). EGFR inhibition is one of the mainstays in the therapy of metastatic CRC[27]. Unfortunately, about 30-50% of CRC patients have a mutated *KRAS* gene that renders them resistant to EGFR targeted therapies [28]. In addition, approximately half of the patients with wild type *KRAS* will develop resistance to EGFR inhibition [29]. Several studies indicate that the extensive cross-talk between EGFR and IGF1R might contribute to the acquired resistance to EGFR inhibition [30].

Therefore, we used the above computational pipeline to analyse a publicly available dataset containing perturbation responses of EGFR and IGF1R STNs in six genetically diverse CRC cell lines, namely HCT116, HT-29, LIM1215, SW403, SW480 and RKO.[20] This dataset comprises measurements of phosphorylation changes in eight key signalling proteins (AKT^{S473}, $ERK2^{T185/Y187}$, $MEK1^{S217/S221}$, $p70S6K^{T421/S424}$, $IGF\text{-}1R^{Y1131}$, $GSK3\alpha/\beta^{S21/S9}$, $IkB\alpha^{S32/S36}$, and $IRS\text{-}1^{S636/S639}$) in response to stimulation with TGFα (a ligand

of EGFR) and IGF-1, and four different inhibitory drugs, respectively (Figs. 2A). As RKO cells hardly showed any drug responses, they were not considered in our analysis.

Firstly, we developed a literature-based generic STN for EGFR and IGF1R pathways consisting of all known interactions (Supp. Note 1) between the measured proteins across a diverse range of cell types. The generic STN represents our prior knowledge of EGFR and IGF1R pathways and was used to formulate the prior distribution for the BMRA algorithm (M&M). BMRA with the above prior was used to reconstruct separate models for TGFα and IGF-1 stimulated STNs in five CRC cells from the corresponding perturbation datasets. This resulted in ten network models (Figs. 2B-E) corresponding to five CRC cell lines, stimulated by two ligands (TGFα or IGF-1). In the reconstructed networks, each interaction is represented by the probability distribution of its strength and the probability of its existence (Figs. 2B,D). The reconstructed models revealed several differences in the signal integration, crosstalk and feedback mechanisms of the EGFR and IGF1R pathways across different CRC cell lines. For instance, ERK crosstalks with the IGF1R pathway by inhibiting AKT phosphorylation in LIM1215 and SW403 cells, but this inhibition is much weaker in other cell lines especially in the IGF-1 stimulated networks. Vice versa, AKT crosstalks with the EGFR pathway by inhibiting MEK phosphorylation in IGF-1 stimulated SW480 cells, but this cross-talk is weaker or non-existent in other cell lines. Main differences in signal integration affect the phosphorylation of GSK3β and p70S6K by ERK and AKT. GSK3β is co-regulated by both ERK and AKT in HCT116 cells, but mostly by ERK alone in HT29 and LIM1215 cells, and by AKT in SW403 cells. Similarly, p70S6K phosphorylation is regulated by ERK in all cells lines, and co-regulated by AKT in all but SW403 cells. Feedback loops were also predicted to operate differentially between cell lines. For instance, p70S6K forms a negative feedback loop by inhibiting IRS1 by phosphorylating its inhibitory sites (S636/639) in TGFα stimulated HCT116 and SW480 cells, but not in other cell lines. Interestingly, both

HCT116 & SW480 are known to be resistant to EGFR inhibitor drugs [31] and it is well know that negative feedback loops may play important roles in drug resistance [32]. Therefore, we chose this feedback loop for further exploration and performed biochemical experiments to (a) validate its existence (b) determine whether its existence is associated with any of the commonly occurring oncogenic mutations.

KRAS mutation is necessary but may not be sufficient for the feedback inhibition of IRS1 by p70S6K

We chose three cell lines for the biochemical experiments, (a) HCT116 cells which has an oncogenic KRAS mutation; (b) HKE3 cells which are genetically identical to HCT116 with the exception that the mutant KRAS allele was knocked out leaving it with a wild-type KRAS allele; and (c) HT29 cells which have a BRAFV600E mutation [33]. We investigated the strength of the p70S6K, IRS1 mediated feedback loop by inhibiting p70S6K phosphorylation and subsequently measuring $IRS1^{S636/S639}$ phosphorylation at 0, 10, 60 minutes after TGFα stimulation. Two different methods were used to inhibit p70S6K phosphorylation, i.e. using small molecule inhibitors of AKT kinase which directly phosphorylates p70S6K, and using small interfering RNA (siRNA) molecules that inhibit p70S6K translation. The AKT inhibitor strongly inhibited p70S6K phosphorylation in all three cell lines (Fig. 3A-C), but significantly reduced $IRS1^{S636/S639}$ phosphorylation only in HCT116 cell line. The siRNA mediated p70S6K knockdown caused a significant reduction in $IRS1^{S636/S639}$ phosphorylation in HCT116 cells at 60 minutes after TGFα stimulation (Fig 3D), but increased $IRS1^{S636/S639}$ phosphorylation in HKE3 and HT29 cell lines (Fig. 3E,F). These observations suggest that the p70S6K and IRS1 mediated negative feedback is prominent in HCT116 cells but it is weaker or non-existent in HT29 and HKE3. The differences in this feedback between HCT116 and HT29 cells is in agreement with the results of the BMRA algorithm.

Additionally, the differences in this feedback between HCT116 and HKE3 cells which differ only by a KRAS mutation suggests that KRAS mutation plays a role in the functioning of this feedback. However, the BMRA algorithm did not detect this feedback in SW403 cells which also have the KRAS mutation. This may be because (a) noise in the perturbation dataset prevented the BMRA algorithm from detecting this feedback in SW403 cells, or (b) KRAS mutation alone is not sufficient for the functioning of this feedback loop, as it may depend on other mutations in HCT116 cells which are absent in SW403 cells. A detailed mutational association analysis of the above feedback loop is out of the scope of this article. However, from the above analysis it is safe to say that KRAS mutation is necessary but may not be sufficient for the operation of the P70S6K, IRS1 mediated negative feedback loop.

A summary of the above results are shown in Fig. 3G. Interestingly, both HCT116 & SW480 which contain the feedback loop in question are known to be resistant to EGFR inhibitor drugs [31], begging the question whether it plays any role in their drug resistance mechanisms. We developed a predictive model of the EGFR & IGF1R pathway in HCT116 cells using the BMM algorithm to answer this question.

Mechanistic modelling of EGF & IGF-1 STN in HCT116 cells using the BMM algorithm

First, we developed a simplified ODE model of the generic EGFR & IGF1R pathways containing their core elements, based on existing literature[10, 13-16, 23, 24] (Fig. 4A). Briefly, the model features the activation of the EGFR and IGF1R by their ligands EGF or TGFα and IGF-1 or Insulin, respectively. The activated receptors induce tyrosine phosphorylation of IRS-1, a multi-adaptor protein that activates RAS and PI3K signalling. The EGFR also can activate RAS independently of IRS-1 leading to activation of the RAF-MEK-ERK kinase cascade. PI3K induces the activation of AKT, which together with ERK phosphorylates and

activates p70S6K. The model also contains crosstalk circuits between ERK and AKT, and AKT and RAF, as well as several feedback loops, e.g. the inhibitory feedback phosphorylation of RAF, MEK, IRS1 by ERK and inhibitory feedback phosphorylation of IRS1 by p70S6K. Although protein activation/deactivation/inhibition are multistage processes involving multi-site phosphorylations, conformation changes, dimerization etc., we simplified these processes to single partial steps[10] that transform a protein from its inactive to active/inhibited form or vice versa. These regulations were modelled using Michaelis-Menten kinetics, except the inhibition of AKT phosphorylation by ERK, which was modelled by a hyperbolic modifier function.[34] It should be noted that EGFR, RAS, RAF, PI3K were not part of reconstructed STN models, as they were not experimentally measured in the perturbation experiment. However, these were included in the ODE model to incorporate the effects of KRAS and PI3K mutations that occur in the HCT116 cells in the ODE model. GSK3β and IkB were part of the reconstructed STNs, but were omitted from the ODE model as these are not involved in the negative feedback regulation of IRS-1 by p70S6K. The resulting model has 66 unknown parameters. The BMM algorithm identified 17 parameters (highlighted yellow in Fig. 4A) which have the maximum impact on the model output (Supp. Note 3) and estimated the HCT116 specific posterior distributions of these parameters (Supp. Note 3), while keeping the remaining parameters at fixed values (Supp. Note 3). Interestingly, the posterior distributions (Fig. 4B) reflect well-known characteristics of EGFR and IGF1R pathways in HCT116 cells. For instance, the activating KRAS mutation in HCT116 cells was well described by the posterior distributions of the RAS (in)activation rates, which indicated that RAS is activated at a much higher rate ($k_{ras1} = 1.35 \pm 0.36\ min^{-1}$, $k_{ras2} = 3.02 \pm 1.5\ min^{-1}$, see Fig. 4B) than it is inactivated ($V_{M_{ras}} = 0.012 \pm 0.18\ a.u.\ min^{-1}$). Additionally, the inferred posterior distributions of many model parameters are bimodal (Fig. 4B), reflecting the true behaviour of the EGF and IGF-1 pathway components since these are

known to be activated or inactivated at different rates depending on whether the pathways are triggered by EGF/TGFα or IGF-1/Insulin.[1, 10, 16] .

Initial validation of the mechanistic model

Model simulation was performed using an ensemble of ODE models, each of which was fitted with parameter values sampled from the corresponding posterior distribution (Supp. Note. 3). The average and standard deviation of the ensemble simulations were then used to represent simulation results and the corresponding confidence intervals. Firstly, we verified whether the interaction strengths simulated by the calibrated model resembled those estimated by the BMRA algorithm. This was indeed the case (Supp. Fig. S1). Subsequently, we simulated ERK and AKT activation levels for different levels of EGF/TGFα or IGF-1/Insulin. The model simulations suggested that ERK and AKT were transiently activated at low levels of EGF and Insulin stimulation, but then quickly achieved sustained activation at higher levels (Fig. 5A-D). These were then validated in biochemical experiments (Fig. 5E-G) confirming the validity of the model. Hence, we used this model to explore the role of p70S6K and IRS1 mediated negative feedback loop in the EGFR inhibitor resistance of HCT116 cells.

Knocking down p70S6K sensitizes HCT116 cells to EGFR inhibitors *in silico*, *in vitro* and *in vivo*

Model simulations indicated that knocking down p70S6K has no effect on ERK phosphorylation, but increases AKT phosphorylation (Fig 6A,B) in EGF/TGFα stimulated HCT116 cells. These simulations were subsequently confirmed in biochemical experiments

(Figs. 6C,D). These results suggest that the p70S6K and IRS1 mediated feedback loop has a significant impact on AKT phosphorylation, but not ERK phosphorylation. Hence, we focussed our attention on AKT activity in serum grown HCT116 cells which approximates the behaviour of cells growing in their natural environment. First, we simulated active AKT concentration in response to separate treatments by p70S6K siRNA and EGFR inhibitors. As in the case of EGF/TGFα stimulated cells, p70S6K knockdown led to an increase of active AKT levels (Fig. 6E) in serum grown cells, whereas EGFR inhibition was rather ineffective on AKT activation (Fig. 6E). These simulations were then confirmed experimentally (Fig. 6F,G). We then simulated active AKT levels in response to combinations of different p70S6K knockdown efficiencies and EGFR inhibitor strengths (Supp. Note 3) in serum grown HCT116 cells. The simulation suggested that a combination of partial p70S6K knockdown and EGFR inhibitor results in significant reduction in AKT phosphorylation compared to either individual treatments (Fig. 6H). This was confirmed in biochemical experiments (Fig. 6I). A possible reason behind the counterintuitive behaviour of AKT in response to p70S6K knockdown is as follows. AKT activity is tightly controlled by several negative and positive feedback loops (Supp. Fig S6) and some of these (Fig. S6 D, E, F) involve p70S6K. The negative feedback (Fig. S6F) involving AKT and p70S6K is usually operational and causes an increase in AKT phosphorylation in response to p70S6K knockdown (Fig 6A,B,C,E,F,G). However, the positive feedback loops (Fig. S6 D,E) involve the RAF-MEK-ERK pathway which becomes saturated in presence of EGF/TGFα, aided by mutant RAS, and hence does not respond to p70S6K knockdown (Fig. 6 A,B,D, S7). EGFR inhibitors releases this pathway from such saturating effect and allow the positive feedbacks involving p70S6K and AKT to operate. When both positive and negative feedbacks involving p70S6K and AKT are operational, AKT phosphorylation may either increase or decrease in response p70S6K knockdown depending on which feedback has more impact on its

phosphorylation. The positive feedbacks involving p70S6K are reinforced by several other positive feedback loops (Fig. S6 A,B,C) involving the RAF-MEK-ERK pathway. Therefore, it is possible that, in presence of EGFR inhibitors, the positive feedbacks have higher impact on AKT phosphorylation than the negative feedbacks leading to a decrease in AKT phosphorylation in response to p70S6K knockdown.

Inhibition of AKT pathway is known to increase apoptosis and decreases invasiveness of some cancer cells.[35-40] Therefore, we examined whether the same is true for HCT116 cells. We tested the effects of combining EGFR inhibition with p70S6K knockdown on serum grown HCT116 cells (Fig. 6F-M). The combination induced significantly higher level of apoptosis in HCT116 cells compared to either individual treatment (Fig. 6J). The Bliss independence score (BIS) of the combination, i.e. the $\log_2$ ratio of the expected to observed cell mortality rates, was found to be -0.217 suggesting that the combined treatment killed more cells than expected, and therefore had a synergistic effect on HCT116 cell death. Furthermore, silencing of p70S6K followed by Lapatinib (an EGFR inhibitor) treatment[1] reduced the invasive potential of the xenografted HCT116 cells more effectively than Lapatinib treatment alone, in *in vivo* invasion assays in zebrafish (*Danio rerio*) (Fig. 6 K,L,M).

These results suggest that the negative feedback loop from p70S6K to IRS1 plays a role in the EGFR inhibitor resistance of the HCT116 cell line and inhibition of p70S6K could overcome such resistance. Interestingly, HKE3 and HT29 cell lines, which do not have this feedback loop, were responsive to EGFR inhibition and were not affected by the p70S6K knockdown (Supp. Figs S9,S10). This further strengthens the hypothesis that p70S6K

[1] BIBX1382 is highly toxic to Zebrafish embryos, hence we used Lapatinib for this experient

mediated negative feedback loop plays a role in the EGFR inhibitor resistance of HCT116 cells.

Discussion

Here, we presented a computational prototype that integrates top-down and bottom-up systems biology approaches in one platform. It allows both reconstructing and mechanistic modelling of STNs from noisy and incomplete perturbation datasets in an integrated analysis pipeline. The resulting framework overcomes some of the most restricting limitations of other similar algorithms (e.g. MRA), such as sensitivity to noise [8], large experimental requirements [8], inability to predict the effects of sequential perturbations etc. We demonstrated this by successfully predicting the effects of individual and combined treatments using p70S6K knockdown and EGFR inhibitors on HCT116 cells using a small set of highly incomplete and noisy data. Therefore, our algorithm pinpointed a potentially novel therapeutic approach to overcome EGFR inhibitor resistance in colorectal cancer patients.

Nevertheless, our algorithm is not without limitations. Firstly, the ODE model was calibrated to exhibit similar local interaction strengths as the probabilistic STNs, which were estimated from steady state perturbation response data. Therefore, these models are likely to correctly simulate steady state behaviour, but may not accurately simulate pathway dynamics. This could be remedied by employing additional time course data for calibration. Secondly, our algorithm is computation intensive and in its current form can be used to analyse networks containing up to 25 nodes within reasonable time. For bigger networks the number of parameters that needs to be calibrated increases dramatically, which would require employing specialised MCMC methods designed for exploring high-dimensional parameter spaces, e.g. Riemann Manifold Langevin adjusted Metropolis Hastings methods.[26] However,

in the current form it is a useful tool for exploring the mechanistic roles of STNs in disease and drug resistance and predicting new therapeutic targets.

Materials and Methods:

The BMRA algorithm:

Let us consider an *N* node ($n_i, i = 1 \dots N$) biochemical network. Modular Response Analysis [17] showed that following a perturbation (p_k), the fractional change ($R_{ik} = \frac{dlog(x_i^{SS})}{dp_k}$) in the steady state concentration (x_i^{SS}) of a network node (n_i) is linearly related to that ($R_{jk}, j \neq i$) of the other nodes ($n_j, j \neq i$) via the corresponding interaction strengths ($r_{ij}, j \neq i$).

$$\sum_{j=1, j\neq i}^{n} r_{ij} R_{jk} = R_{ik}; i \neq k;\ i, k = 1 \dots n \qquad (1)$$

If node $n_j, j \neq i$ directly influences node n_i, then $r_{ij} \neq 0$, otherwise $r_{ij} = 0$. We modified this equation by introducing two more variables, a binary variable A_{ij} in Eq. 1 to indicate whether n_j directly influences n_i ($A_{ij} = 1$) or not ($A_{ij} = 0$), and an error variable (ϵ_{ik}) to account for the imbalance in Eq. 1 caused by noise experimental data. The resulting modified MRA equation is shown below.

$$\sum_{j=1, j\neq i}^{n} A_{ij} r_{ij} R_{jk} + \epsilon_{ik} = R_{ik}; i \neq k;\ k = 1 \dots N \qquad (3)$$

We used Bayesian statistics to infer the probability distributions of $A_{ij}, i \neq j$ and $r_{ij}, i \neq j$ from experimentally measured $\boldsymbol{R}$. First, we conceptually divided the N node network into N smaller sub-networks, each of which consists of a specific node (n_i) and its regulators. The probability distributions of the network descriptors ($\boldsymbol{A_i} = \{A_{ij}, j \neq i\}, \boldsymbol{r_i} =$

$\{r_{ij}, j \neq i\}$) of each of these sub-networks were inferred separately. We first formulated prior probability distributions of $\boldsymbol{A_i}$ and $\boldsymbol{r_i}$ based on existing knowledge of the network topology. Let the binary vector $\boldsymbol{\mathcal{A}_i}$ denote the known regulators of node n_i **.** The prior distribution ($\boldsymbol{P}(\boldsymbol{A_i})$) of $\boldsymbol{A_i}$ was then formulated as $P(\boldsymbol{A_i}) \propto \boldsymbol{exp}(-\psi \times \boldsymbol{d_H}(\boldsymbol{A_i}, \boldsymbol{\mathcal{A}_i}))$, where ψ is a constant and $\boldsymbol{d_H}(\boldsymbol{A_i}, \boldsymbol{\mathcal{A}_i})$ is the Hamming distance between $\boldsymbol{A_i}$ and $\boldsymbol{\mathcal{A}_i}$. $P(\boldsymbol{A_i})$ favours sub-networks which are in agreement with existing knowledge and penalize those which are in disagreement. The extent of penalty depends on ψ with larger values resulting in stiffer penalties. The prior distribution of $\boldsymbol{r_i}$ is dependent on $\boldsymbol{A_i}$ and is denoted by $P(\boldsymbol{r_i}|\boldsymbol{A_i})$. In absence of a direct interaction from n_j to n_i (i.e. $A_{ij} = 0$) the corresponding interaction strength (r_{ij}) was assumed to have 0 value with probability 1, whereas the interaction strengths representing direct interactions ($A_{ij} = 1, j \neq i$) were assumed to have Gaussian priors (Supp. Note 2, [8, 41]). The above prior distributions ($P(\boldsymbol{A_i}), P(\boldsymbol{r_i}|\boldsymbol{A_i})$) were then used to calculate the likelihood ($P(\boldsymbol{R}|\boldsymbol{r_i}, \boldsymbol{A_i})$) of the global changes ($\boldsymbol{R}$) given the network descriptors ($\boldsymbol{A_i}$ **,** $\boldsymbol{r_i}$) (Supp. Note 2 and [8]). Subsequently, we applied Bayes' rule [25] to calculate the joint posterior distribution ($P(\boldsymbol{A_i}, \boldsymbol{r_i}\,|\boldsymbol{R})$) for the network descriptors ($\boldsymbol{r_i}$, $\boldsymbol{A_i}$). Using chain rule, the above joint posterior ($P(\boldsymbol{A_i}, \boldsymbol{r_i}\,|\boldsymbol{R})$) is the product of the marginal posterior ($P(\boldsymbol{A_i}, |\boldsymbol{R})$) of the binary indicators ($\boldsymbol{A_i}$) and the conditional posterior ($P(\boldsymbol{r_i}\,|\boldsymbol{A_i}, \boldsymbol{R})$) of the interactions strengths ($\boldsymbol{r_i}$), i.e. $P(\boldsymbol{A_i}, \boldsymbol{r_i}\,|\boldsymbol{R}) = P(\boldsymbol{r_i}|\boldsymbol{A_i}, \boldsymbol{R}) \times P(\boldsymbol{A_i}|\boldsymbol{R})$. The marginal posterior ($P(\boldsymbol{A_i}, |\boldsymbol{R})$) is proportional to the product of a multivariate Student t (MVSt) and the prior distribution, whereas the conditional posterior ($P(\boldsymbol{r_i}|\boldsymbol{A_i}, \boldsymbol{R})$) is proportional to a MVSt distribution (Supp. Note 2). The constant of proportionalities of the above posterior could not be calculated analytically, and therefore we used an MCMC algorithm to approximate the true joint posterior $P(\boldsymbol{A_i}, \boldsymbol{r_i}\,|\boldsymbol{R})$. The MCMC algorithm starts with a random network, and in each iteration (t) randomly proposes a new sub-network ($\boldsymbol{A_i^{tn}}$) of $\boldsymbol{\mathcal{A}_i}$ which differs from the current sub-network by only one interaction (see Supp. Note) [5, 42]. The proposed sub-network

was then accepted ($\boldsymbol{A_i^t = A_i^{tn}}$) with probability $\alpha = \min\left(1, \frac{P(\boldsymbol{A_i^{tn}}|\boldsymbol{R})}{P(\boldsymbol{A_i^t}|\boldsymbol{R})}\right)$ [43](Supp. Note 2). Subsequently, we drew a sample ($\boldsymbol{r_i^t}$) from the conditional distributions $P(\boldsymbol{r_i}|\boldsymbol{A_i^t}, \boldsymbol{R})$ using a random vector generator that samples from custom MVSt distributions. The above two steps were repeated for a large number (10^5) of iterations and after an initial burn-in period (5×10^4 iterations), the accepted network topologies and interaction strengths were recorded. The mean of an element A_{ij} of $\boldsymbol{A_i}$ over the post-burn-in samples was then used as the posterior probability of the interaction from node n_j to n_i, i.e. $P(A_{ij}|\boldsymbol{R}) = \frac{1}{5\times10^4}\sum_{t=5\times10^4+1}^{10^5} A_{ij}^t$. On the other hand, the sample means/variance/covariance of different interaction strengths (r_{ij}) across the post-burn-in samples were used as the posterior mean/variance/covariance of the strengths of the interaction from node n_j to n_i,. See Supp. Note 2 for further details of implementation of the above algorithm. A MATLAB implementation of the above (BMRA) algorithm is freely available from http://www.ucd.ie/sbi/.

Calibrating the parameters of the ODE model

To calibrate the parameters ($\boldsymbol{\Theta}$) of the ODE model, we first identified those parameters ($\boldsymbol{\Theta_s} \subseteq \boldsymbol{\Theta}$) which had the maximum impact on the model output(s). Here, the model output (M_o) is defined as the area under the concentration vs time curve ($x_o(t)$) of the output node (n_o) of the model ,i.e. $M_o = \int_{t=0}^{\infty} x_o(t)dt$. When a model has many output nodes (n_{o1}, n_{o2} ...) then the sum of the areas of their concentration curves ($x_{o1}(t), x_{o2}(t), ..$) is used as model output ($M_o = \int_{t=0}^{\infty}(x_{o1}(t) + x_{o2}(t) + ..)dt$). If a small change ($\Delta\theta_k$) in a parameter ($\theta_k$) causes the model output to change by ΔM_o, then the sensitivity (S_k) of the model to this parameter is given by $S_k = \log\left(\left|\frac{\Delta M_o}{\Delta\theta_k}\right|\right)$. These sensitivities quantify the impact of a model's

parameters on its output. Model sensitivities (S_k) were estimated for each model parameter (θ_k) for a large number (10^5) of randomly generated (see Supp. Note 3) parameter settings using Monte Carlo sampling (Supp. Fig. SN3.1). The parameters which had low (≈ 0) average sensitivities were assigned fixed values (Supp. Note 3) and those with higher average sensitivities ($\gg 0$) were calibrated using a Bayesian parameter estimation algorithm. First, we assigned a prior probability distribution ($P(\boldsymbol{\Theta_s})$, Supp. Note 3) to the sensitive parameters ($\boldsymbol{\Theta_s}$) and formulated a likelihood function ($\mathcal{L}(\boldsymbol{\Theta_s})$) for the same: $\mathcal{L}(\boldsymbol{\Theta_s}) \propto \prod_l |\boldsymbol{\Sigma_{rl}}|^{-\frac{1}{2}} \boldsymbol{exp}\left(-(\boldsymbol{r_{Ml}} - \boldsymbol{\mu_{rl}})^T \boldsymbol{\Sigma_{rl}^{-1}} (\boldsymbol{r_{Ml}} - \boldsymbol{\mu_{rl}})\right)$. Here, $\boldsymbol{r_l}$ represents the simulated interaction strengths in condition (e.g. ligand stimulation) l; $\boldsymbol{\mu_{rl}}$ and $\boldsymbol{\Sigma_{rl}}$, are their means and covariance matrices as estimated by the BMRA algorithm from experimental data. We then used Bayes' rule [25] to calculate the posterior distribution ($P(\boldsymbol{\Theta_s}|\boldsymbol{\mu_{rl}}, \boldsymbol{\Sigma_{rl}}, l = 1,2 \ldots.)$) of the sensitive parameters ($\boldsymbol{\Theta_s}$): $P(\boldsymbol{\Theta_s}|\boldsymbol{\mu_{rl}}, \boldsymbol{\Sigma_{rl}}, l = 1,2 \ldots.) \propto \mathcal{L}(\boldsymbol{\Theta_s}) P(\boldsymbol{\Theta_s})$. However, the posterior distribution could only be calculated up to a constant of proportionality. Therefore, MCMC sampling was used to estimate the true posteriors of these parameters. The MCMC algorithm starts with a random set of values for the sensitive parameters, and in each iteration (t) proposes a new set of values ($\boldsymbol{\Theta_s^{tn}}$) based on their current values ($\boldsymbol{\Theta_s^t}$) using a Gaussian proposal distribution ($q(\boldsymbol{\Theta_s^{tn}}|\boldsymbol{\Theta_s^t})$) which is defined in log-space, i.e. $\log(\boldsymbol{\Theta_s^{tn}}) \sim \mathcal{N}\left(\log(\boldsymbol{\Theta_s^t}), \boldsymbol{V_{\Theta_s}}\right)$, where $\boldsymbol{V_{\Theta_s}}$ is the covariance matric of the proposal distribution (see Supp. Note 3) . The new values ($\boldsymbol{\Theta_s^{tn}}$) are then accepted ($\boldsymbol{\Theta_s^t} = \boldsymbol{\Theta}\ \boldsymbol{_s^{tn}}$) with probability $\alpha = \min\left(1, \frac{P(\boldsymbol{\Theta_s^{tn}}|\boldsymbol{\mu_{rl}}, \boldsymbol{\Sigma_{rl}}, l = 1,2 \ldots.)}{P(\boldsymbol{\Theta_s^t}|\boldsymbol{\mu_{rl}}, \boldsymbol{\Sigma_{rl}}, l = 1,2 \ldots.)}\right)$ [43](Supp. Note 3). This process was repeated a large number of times (2.5×10^5) and the samples drawn after a burn-in (10^5) period were used to estimate the posterior distribution of $\boldsymbol{\Theta_s}$. The variances ($\boldsymbol{V_{\Theta_s}}$) of the proposal distribution were adaptively tuned to improve mixing of the MCMC chains [44]. For a robust exploration of the parameter space, we deployed multiple parallel Markov Chains which operate at

different “temperatures” (parallel tempering) [45]. The chains which operate at high temperatures (temperature>>1) take large steps and can explore a vast region in the parameter space relatively quickly. On the other hand, the chains which operate at low temperature (temperature $\approx$ 1) take small steps and perform local searches. These chains pass information among each other in a probabilistic manner helping the Markov chains operating at low temperature to avoid getting stuck at local minima. The samples drawn by the chain operating at normal temperature (temperature$=$ 1) were used to estimate the posterior distribution of $\boldsymbol{\Theta_s}$. A MATLAB implementation of the above calibration algorithm is freely available from http://www.ucd.ie/sbi/.

Cell lines

The human colorectal adenocarcinoma cell line HT29 was acquired from the European Collection of Cell Cultures (ECACC, Salisbury, UK). The human colorectal carcinoma cell lines HCT116 and HKE-3 were kindly provided by Dr Senji Shirasawa.[45] Cells were maintained in DMEM (Dulbecco's modified Eagle's medium; Invitrogen Life Technologies, Paisley, UK) supplemented with 10% foetal bovine serum (Invitrogen Life Technologies, Paisley, UK), 100 IU/ml penicillin and 100 μg/ml streptomycin (Invitrogen Life Technologies, Grand Island, NY, USA); and incubated at 37°C in a humidified atmosphere of 5% CO2.

Reagents

The following inhibitors were used: PI-3K inhibitor LY294002 (10 μM, Calbiochem, La Jolla,CA, USA), Akt inhibitor IV (10 μM, EMD Millipore, Billerica, MA), Akt inhibitor VIII (10 μM, Sigma-Aldrich, St Louis, MO, USA), Rapamycin (250 nM, Calbiochem, San

Diego, CA, USA), MEK inhibitor U0126 (10 μM, Promega, Madison, WI, USA), EGFR inhibitor BibX 1382 (10 μM, Calbiochem, San Diego, CA, USA) and Lapatinib (2.5 μM, Selleck Chemicals). Cells were treated with 100 ng/mL TGF-alpha (Peprotech, Rocky Hill, NJ, USA), 100 ng/mL EGF (Peprotech, Rocky Hill, NJ, USA) or 10 μg/mL insulin (Sigma-Aldrich, St Louis, MO, USA).

RNA interference

Validated siRNA to interfere with RPS6KB1 mRNA (Sense: 5'-> GGUUUUUCAAGUACGAAAAtt Antisense: UUUUCGUACUUGAAAAACCtt) and non-targeting control siRNA were pre-designed by Ambion (Austin, TX, USA). Transfections were performed with Oligofectamine (Invitrogen Life Technologies, Carlsbad, CA, USA) according to the manufacturer's instructions. Briefly, diluted oligonucleotides were pre-incubated with Oligofectamine for 20 min at room temperature then the mixtures were added dropwise to 1.5×10^5 cells in OptiMEM (Invitrogen Life Technologies, Paisley, UK). After 4 h at 37 °C, DMEM with 30 % FBS was added to the cultures. Cells were lysed after 24, 48 or 72 h for Western blot analyses to test the efficiency of knock down.

Quantitative immunoblotting

Cells were serum-starved for 4 hours prior to stimulation. Cells were lysed in 1x Lysis Buffer (Cell Signaling Technology Inc., Danvers, MA, USA) according to the manufacturer's instructions. NuPAGE Bis-Tris gels (Novex, Life Technologies, Carlsbad, CA) were used to resolve proteins. Gels were transferred to PVDF membrane at 30 V for 1h. Membranes were blocked with TBS-Tween (pH 7.4) containing 5 % skim milk for an hour and incubated with 1:1000 diluted rabbit anti-human phospho-Akt (Ser473), Akt, phospho-p44/42 MAPK (Thr202/Tyr204), p44/42 MAPK, p70 S6 Kinase, phospho-p70 S6 Kinase (Thr389), IRS-1,

phospho-IRS-1 (Ser636/639) antibodies or as control with mouse anti-human GAPDH (all from Cell Signaling Technology Inc., Danvers, MA, USA) in TBS-Tween with 5 % BSA overnight at 4 °C. After washing, bound antibodies were detected with 1:5000 diluted horseradish peroxidase-conjugated anti-rabbit or anti-mouse IgG (Cell Signaling Technology Inc., Danvers, MA, USA), followed by development with WesternBright chemiluminescent substrate (Advansta, Menlo Park, CA, USA). Membranes were scanned by using a high-sensitivity CCD camera (Chemi Image, Advanced Molecular Vision, Grantham, UK). Quantification of the bands was performed by using the ImageJ software.

Apoptosis assays

Annexin V staining: The Annexin-V-FITC Apoptosis Detection Kit (BD Biosciences, San Diego, CA, USA) was used to detect apoptosis by flow cytometry. Cells were exposed to BibX 1382 treatment for 48 h, then harvested (including detached cells) and processed according to the manufacturer's instructions. Briefly, cells were washed twice with PBS and resuspended in 1X Annexin V Binding Buffer. Then $1x10^5$ cells were incubated with 5 µl Annexin V-FITC and 5 µl propidium-iodide for 15 min at room temperature. Finally, 400 µl of 1x Binding Buffer was added. Labelled cells were analysed with an Accuri C6 flow cytometer (Beckton Dickinson Immunocytometry Systems) equipped with the Accuri C6 software program (Becton Dickinson) for data acquisition and analysis.

Sub-G_1 analysis. Alternatively, following the drug treatments, cells were fixed in ethanol and stained with propidium-iodide. The sub-G1 population was determined using an Accuri C6 flow cytometer.

Zebrafish xenografts

Zebrafish (*Danio rerio*) maintenance and experiments were approved by the UCD Animal Research Ethics Committee. Wild-type (AB line) embryos were produced by natural spawning and kept on a 14h/10h light/dark cycle at 28.5 °C. Control or p70S6K knock-down HCT116 cells were labelled with DiO or DiI (Vybrant; Invitrogen Life Technologies, Carlsbad, CA, USA) lipophilic fluorescent dyes, respectively, as described earlier. Equal amounts of control and p70S6K knock-down HCT116 cells were mixed, resuspended in PBS and kept on ice before microinjections. Two days post-fertilization, dechorionated zebrafish embryos were anesthetized by 0.016% MS-222 (Sigma- Aldrich, Steinheim, Germany) and approximately 100 cells were transplanted into the yolk sac of embryos using a microinjector (PV830 Pneumatic Picopump, World Precision Instruments with M-152 manipulator, Narishige) equipped with a glass capillary (World Precision Instruments, FL, USA). Then embryos were incubated for 1 h at 31 °C and checked for cell presence. Fish with fluorescent cells outside the implantation area were excluded from further analysis. All other fish were incubated with DMSO or 2.5 μM Lapatinib (Selleck Chemicals) at 33 °C for up to 3 days. Disseminations of cells were monitored by an Olympus SZX10 fluorescent stereo microscope equipped with an Olympus DP71 camera.

Acknowledgement

We thank Dr. Dirk Fey for stimulating discussions on several aspects of the computational work. We also thank Alfonso Blanco (Conway Core Technologies, University College Dublin) for his help with flow cytometry and Breandan Kennedy (Zebrafish Facility, University College Dublin) for his useful advice on microinjections. This work was supported by the Science Foundation Ireland under Grant No. 06/CE/B1129 and the European Union FP7 Grant No. 259348-2 “ASSET”.

Author Contributions

M.H. designed and performed the biological experiments and wrote the paper; B.N.K designed the study and wrote the paper, W.K. designed the study and wrote the paper, T.S. designed the study, designed and implemented the computational analysis and wrote the paper.

References

1. Aksamitiene, E., Kiyatkin, A. & Kholodenko, B.N. Cross-talk between mitogenic Ras/MAPK and survival PI3K/Akt pathways: a fine balance. *Biochem Soc T* **40**, 139-146 (2012).
2. Kreeger, P.K., Mandhana, R., Alford, S.K., Haigis, K.M. & Lauffenburger, D.A. RAS mutations affect tumor necrosis factor-induced apoptosis in colon carcinoma cells via ERK-modulatory negative and positive feedback circuits along with non-ERK pathway effects. *Cancer research* **69**, 8191-8199 (2009).
3. Morris, M.K., Saez-Rodriguez, J., Clarke, D.C., Sorger, P.K. & Lauffenburger, D.A. Training signaling pathway maps to biochemical data with constrained fuzzy logic: quantitative analysis of liver cell responses to inflammatory stimuli. *PLoS computational biology* **7**, e1001099 (2011).

4. Tentner, A.R. et al. Combined experimental and computational analysis of DNA damage signaling reveals context-dependent roles for Erk in apoptosis and G1/S arrest after genotoxic stress. *Molecular systems biology* **8**, 568 (2012).

5. Mukherjee, S. & Speed, T.P. Network inference using informative priors. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 14313-14318 (2008).

6. Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D.A. & Nolan, G.P. Causal protein-signaling networks derived from multiparameter single-cell data. *Science* **308**, 523-529 (2005).

7. Sahin, O. et al. Modeling ERBB receptor-regulated G1/S transition to find novel targets for de novo trastuzumab resistance. *BMC systems biology* **3**, 1 (2009).

8. Santra, T., Kolch, W. & Kholodenko, B.N. Integrating Bayesian variable selection with Modular Response Analysis to infer biochemical network topology. *BMC systems biology* **7**, 57 (2013).

9. Bertaux, F., Stoma, S., Drasdo, D. & Batt, G. Modeling dynamics of cell-to-cell variability in TRAIL-induced apoptosis explains fractional killing and predicts reversible resistance. *PLoS computational biology* **10**, e1003893 (2014).

10. Borisov, N. et al. Systems-level interactions between insulin–EGF networks amplify mitogenic signaling. *Molecular systems biology* **5**, 256 (2009).

11. Chen, C., Baumann, W.T., Clarke, R. & Tyson, J.J. Modeling the estrogen receptor to growth factor receptor signaling switch in human breast cancer cells. *FEBS letters* **587**, 3327-3334 (2013).

12. Chen, C. et al. Mathematical models of the transitions between endocrine therapy responsive and resistant states in breast cancer. *Journal of The Royal Society Interface* **11**, 20140206 (2014).

13. Fey, D., Croucher, D.R., Kolch, W. & Kholodenko, B.N. Crosstalk and Signaling Switches in Mitogen-Activated Protein Kinase Cascades. *Frontiers in physiology* **3**, 355 (2012).

14. Gambin, A., Charzynska, A., Ellert-Miklaszewska, A. & Rybinski, M. Computational models of the JAK1/2-STAT1 signaling. *Jak-Stat* **2**, e24672 (2013).

15. Kholodenko, B.N., Demin, O.V., Moehren, G. & Hoek, J.B. Quantification of short term signaling by the epidermal growth factor receptor. *The Journal of biological chemistry* **274**, 30169-30181 (1999).

16. Orton, R.J. et al. Computational modelling of cancerous mutations in the EGFR/ERK signalling pathway. *BMC systems biology* **3** (2009).

17. Kholodenko, B.N. et al. Untangling the wires: a strategy to trace functional interactions in signaling and gene networks. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 12841-12846 (2002).

18. Andrec, M., Kholodenko, B.N., Levy, R.M. & Sontag, E. Inference of signaling and gene regulatory networks by steady-state perturbation experiments: structure and accuracy. *Journal of theoretical biology* **232**, 427-441 (2005).

19. Santos, S.D., Verveer, P.J. & Bastiaens, P.I. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nature cell biology* **9**, 324-330 (2007).

20. Klinger, B. et al. Network quantification of EGFR signaling unveils potential for targeted combination therapy. *Molecular systems biology* **9**, 673 (2013).

21. Tomaru, Y. et al. Identification of an inter-transcription factor regulatory network in human hepatoma cells by Matrix RNAi. *Nucleic acids research* **37**, 1049-1060 (2009).

22. Hatzivassiliou, G. et al. Mechanism of MEK inhibition determines efficacy in mutant KRAS- versus BRAF-driven cancers. *Nature* **501**, 232-236 (2013).

23. Kholodenko, B.N., Hoek, J.B., Westerhoff, H.V. & Brown, G.C. Quantification of information transfer via cellular signal transduction pathways. *FEBS letters* **414**, 430-434 (1997).

24. Tyson, J.J. et al. Dynamic modelling of oestrogen signalling and cell fate in breast cancer cells. *Nature reviews. Cancer* **11**, 523-532 (2011).

25. Bayes, M. & Price, M. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions* **53**, 370-418 (1763).

26. Girolami, M. & Calderhead, B. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**, 123-214 (2011).

27. Krasinskas, A.M. EGFR Signaling in Colorectal Carcinoma. *Pathology Research International* **2011**, 932932 (2011).

28. De Roock, W. et al. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol* **11**, 753-762 (2010).

29. Van Emburgh, B.O., Sartore-Bianchi, A., Di Nicolantonio, F., Siena, S. & Bardelli, A. Acquired resistance to EGFR-targeted therapies in colorectal cancer. *Molecular Oncology* **8**, 1084-1094.

30. Singh, I., Amin, H., Rah, B. & Goswami, A. Targeting EGFR and IGF 1R: a promising combination therapy for metastatic cancer. *Front Biosci (Schol Ed)* **5**, 231-246 (2013).

31. Troiani, T. et al. Primary and acquired resistance of colorectal cancer cells to anti-EGFR antibodies converge on MEK/ERK pathway activation and can be overcome by combined MEK/EGFR inhibition. *Clinical Cancer Research* (2014).

32. Rozengurt, E., Soares, H.P. & Sinnet-Smith, J. Suppression of feedback loops mediated by PI3K/mTOR induces multiple overactivation of compensatory pathways: an unintended consequence leading to drug resistance. *Mol Cancer Ther* **13**, 2477-2488 (2014).

33. Shirasawa, S., Furuse, M., Yokoyama, N. & Sasazuki, T. Altered growth of human colon cancer cell lines disrupted at activated Ki-ras. *Science* **260**, 85-88 (1993).

34. Tsyganov, M.A., Kolch, W. & Kholodenko, B.N. The topology design principles that determine the spatiotemporal dynamics of G-protein cascades. *Mol Biosyst* **8**, 730-743 (2012).

35. Dunn, E.F. et al. Dasatinib sensitizes KRAS mutant colorectal tumors to cetuximab. *Oncogene* **30**, 561-574 (2011).

36. Aksamitiene, E., Kholodenko, B.N., Kolch, W., Hoek, J.B. & Kiyatkin, A. PI3K/Akt-sensitive MEK-independent compensatory circuit of ERK activation in ER-positive PI3K-mutant T47D breast cancer cells. *Cell Signal* **22**, 1369-1378 (2010).

37. Ye, Q., Cai, W., Zheng, Y., Evers, B.M. & She, Q.B. ERK and AKT signaling cooperate to translationally regulate survivin expression for metastatic progression of colorectal cancer. *Oncogene* **33**, 1828-1839 (2014).

38. Jiang, Q., Li, T., Liu, D. & Zhang, H. PI3K/Akt pathway involving into apoptosis and invasion in human colon cancer cells LoVo. *Mol Biol Rep* **41**, 3359-3367 (2014).

39. LIU, J.-J. & DUAN, R.-D. LY294002 Enhances Boswellic Acid-induced Apoptosis in Colon Cancer Cells. *Anticancer Research* **29**, 2987-2991 (2009).

40. Ponnurangam, S., Standing, D., Rangarajan, P. & Subramaniam, D. Tandutinib Inhibits the Akt/mTOR Signaling Pathway to Inhibit Colon Cancer Growth. *Molecular Cancer Therapeutics* **12**, 598-609 (2013).

41. Gupta, M. & Ibrahim, J.G. An Information Matrix Prior for Bayesian Analysis in Generalized Linear Models with High Dimensional Data. *Statistica Sinica* **19**, 1641-1663 (2009).

42. Santra, T. A Bayesian Framework that integrates heterogeneous data for inferring gene regulatory networks. *Frontiers in Bioengineering and Biotechnology* **2** (2014).

43. HASTINGS, W.K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97-109 (1970).

44. Haario, H., Saksman, E. & Tamminen, J. An adaptive Metropolis algorithm. *Bernoulli* **7**, 223-242 (2001).

45. Geyer, C.J. in Proceedings of the 23rd Symposium on the Interface, American Statistical Association, New York, 156. (1991).

Figure legends

Figure 1. Schematic outline of the computational analysis pipeline and its components.

(A) Workflow of the integrated analysis platform.

(B) Workflow of the BMRA algorithm which reconstructs network models of STNs from perturbation datasets.

(C) Workflow of the BMM algorithm which reconstructs ODE models of STNs using reconstructed network models.

Figure 2. Network reconstruction of the EGFR and IGF1R pathways in colon cancer cells.

(A) Schematic diagram depicting reconstruction of EGFR and IGF1R pathways in CRC cells using the BMRA algorithm.

(B-E) Reconstructed network models of cells stimulated with TGFα (**B,C**) or IGF-1 (**D,E**), respectively. Panels B,D show the predicted interaction strengths with red bars indicating positive and blue bars negative interaction strengths. X-axis represents interactions that occur in the EGFR and IGF1R pathways; Y-axis represents cell lines; and Z-axis represents interaction strengths. The height of each bar and the associated error-bar represent the absolute mean and standard deviation of the corresponding interaction strength. The width of each bar represents the corresponding interaction probability. The differences in the topologies of TGFα and IGF-1 stimulated pathways in different CRC cell lines are summarized in panels C,E, respectively. The squares and circles associated with each connection represent positive (+ ve) or negative (- ve) interaction strengths, respectively. The size indicates the interaction strength, the cell lines are colour-coded, and the solid and faded colours represent the mean and standard deviations of each interaction strength.

Figure 3. p70S6K mediated feedback inhibition of IRS1 in HCT116, HKE3 & HT29 cells.

(A,B,C) Effects of AKT inhibition on p70S6K$^{T421/S424}$ and IRS1$^{S636/S639}$ phosphorylation in HCT116, HKE3 and HT29 cells, respectively. Starved cells were treated with AKT inhibitor VIII (10μM) for one hour and then stimulated with TGFα (100nM) for 0, 10, 60 minutes. Phospho-AKT^{S473}, phospho-IRS1$^{S636/S639}$, total IRS1, phospho-p70S6K$^{T421/S424}$ and total p70S6K were measured by Western blotting.

(D,E,F) Effects of p70S6K knockdown (KD) on IRS1 phosphorylation in HCT116, HKE3 and HT29 cells, respectively. Cells were transfected with siRNA against p70S6K. 24 hours later

they were serum starved for 4 hours, treated with TGFα and analysed as above. Blots were quantified using ImageJ. The phospho-proteins were normalized to the respective total proteins, and the normalized levels were scaled between 0 and 1 and plotted. Error bars representing standard deviation were derived from 3 (for A-C) or 2 (for D-F) independent experiments (Figs. S2,S3).

(G) Graphical summary of the results. The squares and circles represent positive (+ ve) or negative (- ve) interaction strengths, respectively.

Figure 4. Mechanistic, predictive modelling of EGFR and IGF1R pathways in HCT116 cells.

(A) The mathematical model of the generic EGFR and IGF1R pathways. Each protein has an inactive and active state denoted by the suffix 'a'. Proteins which are phosphorylated at inhibitory sites also have inhibited states denoted by the suffix 'i'. $MM(k_x, K_x)$ represents Michaelis-Menten kinetics with rate and dissociation constants k_x & K_x respectively; $MM1(K_{Mx}, V_{Mx})$ represents Henri-Michaelis-Menten kinetics with Michaelis constant K_{Mx} and maximum velocity constants V_{Mx}; $HM(k_x, K_x, K_{x1}, \beta)$ represents General Hyperbolic Modifier kinetics[29] with rate constant k_x, dissociation constant K_x, activation/inactivation constant K_{x1}, modification constant β. The molecules in red have activating mutations in HCT116 cells. The model was found to be sensitive to the parameters highlighted in yellow.

(B) Posterior distributions of the sensitive parameters. The X-axis of each subplot represents the values of the corresponding parameter and the Y-axis represents the posterior probabilities associated with these values. The distributions of prior values are represented by grey curves.

Figure 5. Biochemical validation of the mechanistic ODE model.

(A-D) Simulated dynamics of ERK and AKT activation for different levels of EGF/TGFα and Insulin/IGF-1 stimulations, respectively. The X-axis represents time in minutes, the Y-axis EGF or Insulin levels in arbitrary units (AU), and the Z-axis the concentrations of activated ERK (A,C) and AKT (B,D) in AU respectively. Solid lines represent average activities and the shaded areas the corresponding standard deviations.

(E,F,G,H) Phosphorylation levels of ERK1/2$^{T202/Y204}$ and AKT^{S473} at 0, 10 and 60 minutes were determined by Western blotting after stimulation of serum starved HCT116 cells with 20ng/mL EGF, 100ng/mL EGF, 1 μg/mL Insulin or 10μg/mL Insulin, respectively. Blots were quantified using ImageJ. The phosphorylation levels of ERK1/2$^{T202/Y204}$ and AKT^{S473} were then normalized to the loading control GAPDH, scaled between 0 &1 and plotted. The data represent 3 independent experiments (Fig. S4), with error bars indicating standard deviation.

Figure 6. The role of the p70S6K mediated negative feedback to IRS1 in EGFR inhibitor resistance of HCT116 cells.

(A,B) Simulated active ERK (blue) and AKT (red) concentrations following p70S6K knockdown and different levels of EGF/TGFα stimulations. Solid lines and shaded areas represent mean and standard deviations.

(C,D) Effects of p70S6K knockdown on AKT^{S473} and ERK1/2$^{T202/Y204}$ phosphorylation. HCT116 cells were transfected with non-targeting siRNAs (CTRL) or siRNAs against p70S6K (KD) and grown in serum (see Fig. S2 for knockdown efficiency) for 24 hours. Then starved (4h) cells were stimulated by TGFα (100nM) for the indicated timepoints. Protein levels were quantified and normalized as above. Data are representative of 2 independent experiments (Fig. S6).

(E) Simulated effects of p70S6K knockdown and EGFR inhibition on AKT activation in serum grown (EGF/TGFα=0.01 AU, IGF/Insulin=0.01 AU) HCT116 cells.

(F,G) Effects of p70S6K knockdown and EGFR inhibition on AKT^{S473} phosphorylation in serum grown HCT116 cells. CTRL and KD HCT116 cells were treated with EGFR inhibitor (BIBX1382,5μM). Phospho-AKT^{S473} levels were measured at the indicated timepoints after treatment, and analysed as above. Data in (F) is representative of two independent experiments (Fig. S5b).

(H) Simulation of AKT activity for different p70S6K knockdown efficiencies and EGFR inhibitor strengths. AKT activity was quantified by scaled (between 0-1) integral of active AKT levels.

(**I**) Phospho-AKT^{S473} levels in CTRL and KD HCT116 cells with and without BIBX treatment. Data represent two independent experiments (Fig.H and Fig. S5b).

(J) Combined effects of p70S6K knockdown and EGFR inhibiton on apoptosis of HCT116 cells. CTRL and KD cells were treated with 5μM BIBX1382 and apoptosis was measured 48 hours later. Fold-changes in apoptosis with respect to CTRL cells are shown. Results represent 4 independent experiments (Fig. S8).

(K-M) The effect of combining EGFR inhibition with p70S6K knockdown on the invasive potential of xenotransplanted HCT116 cells in zebrafish (M&M, Fig S11). **(K)** Evaluation of microinjected zebrafish at 3dpi. **(L)** Microscopic images of disseminated CTRL (green) and KD (red) HCT116 cells in the zebrafish tail after DMSO or Lapatinib treatment. **(M)** Quantitation of disseminated cells within the zebrafish after DMSO or Lapatinib treatment (mean+/- SEM).

Figures

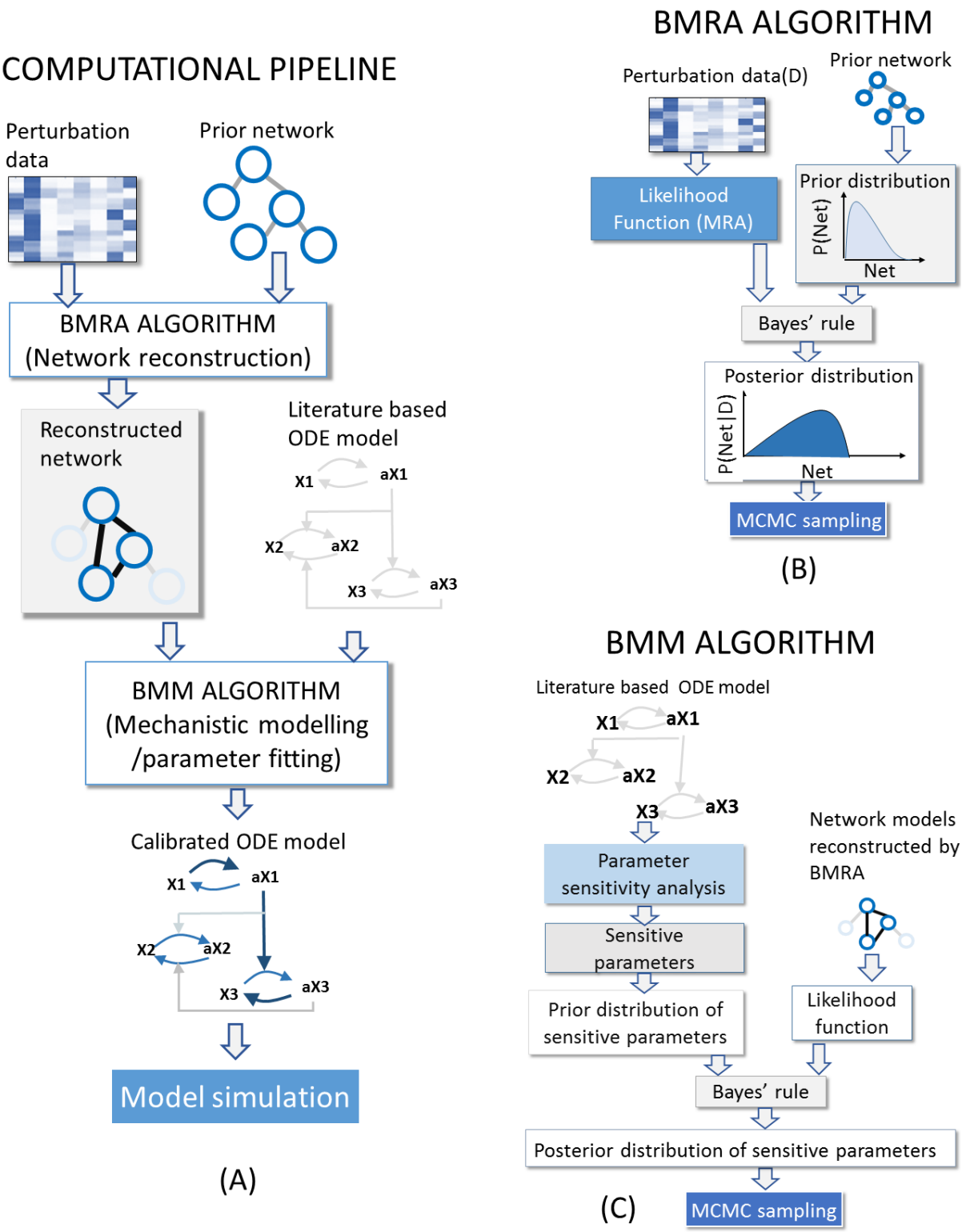
COMPUTATIONAL PIPELINE
Perturbation data
Prior network
BMRA ALGORITHM (Network reconstruction)
Reconstructed network
Literature based ODE model
X1
aX1
X2
aX2
X3
aX3
BMM ALGORITHM (Mechanistic modelling /parameter fitting)
Calibrated ODE model
Model simulation
(A)
BMRA ALGORITHM
Perturbation data(D)
Prior network
Likelihood Function (MRA)
Prior distribution
P(Net)
Net
Bayes' rule
Posterior distribution
P(Net|D)
MCMC sampling
(B)
BMM ALGORITHM
Literature based ODE model
Parameter sensitivity analysis
Sensitive parameters
Prior distribution of sensitive parameters
Network models reconstructed by BMRA
Likelihood function
Bayes' rule
Posterior distribution of sensitive parameters
MCMC sampling
(C)

Figure 1

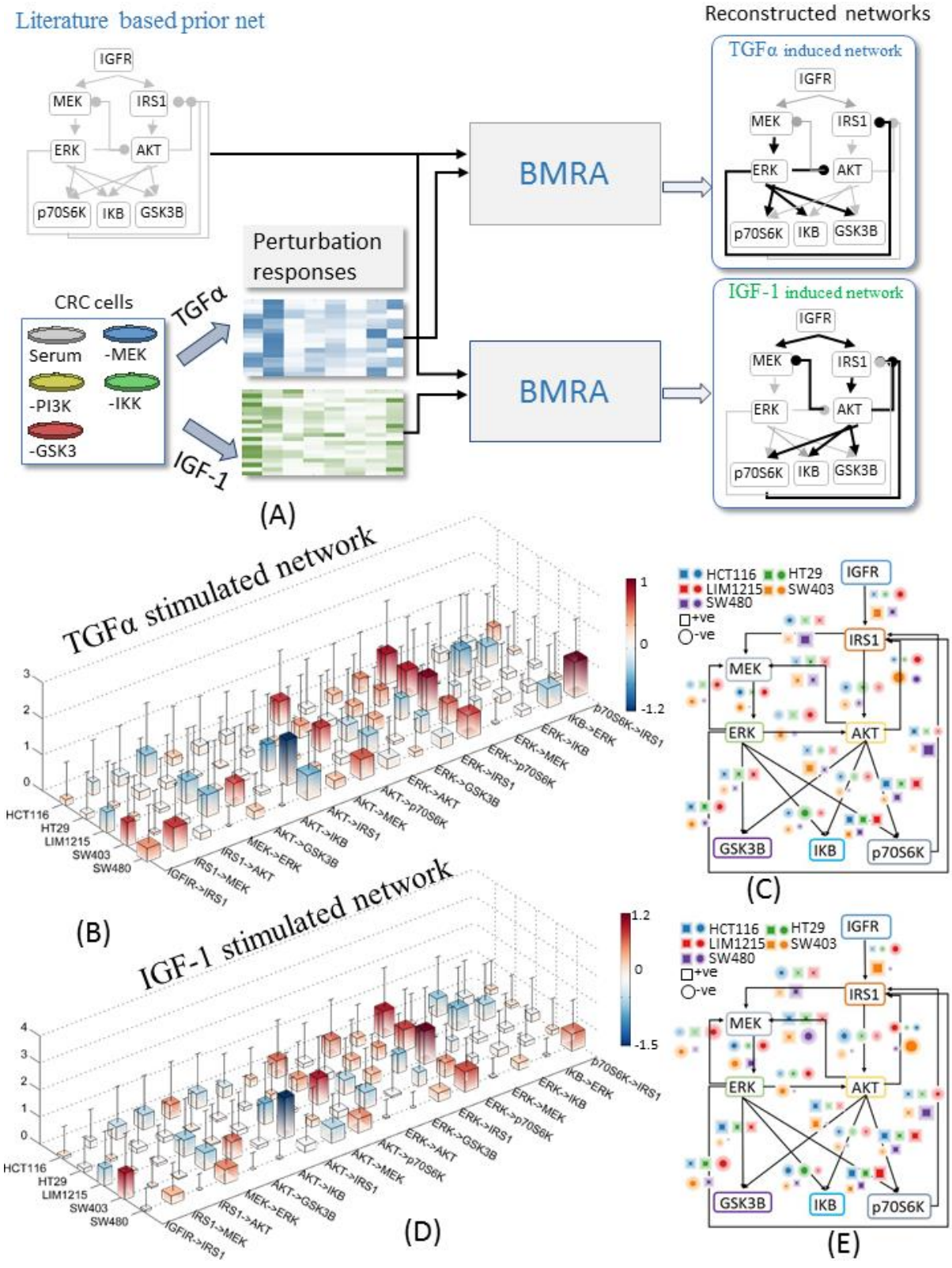
Literature based prior net
IGFR
MEK
IRS1
ERK
AKT
p70S6K
IKB
GSK3B
Reconstructed networks
TGFα induced network
IGF-1 induced network
BMRA
BMRA
Perturbation responses
CRC cells
Serum
-MEK
-PI3K
-IKK
-GSK3
TGFα
IGF-1
(A)
TGFα stimulated network
(B)
IGF-1 stimulated network
(D)
HCT116
HT29
LIM1215
SW403
SW480
+ve
-ve
(C)
(E)

Figure 2

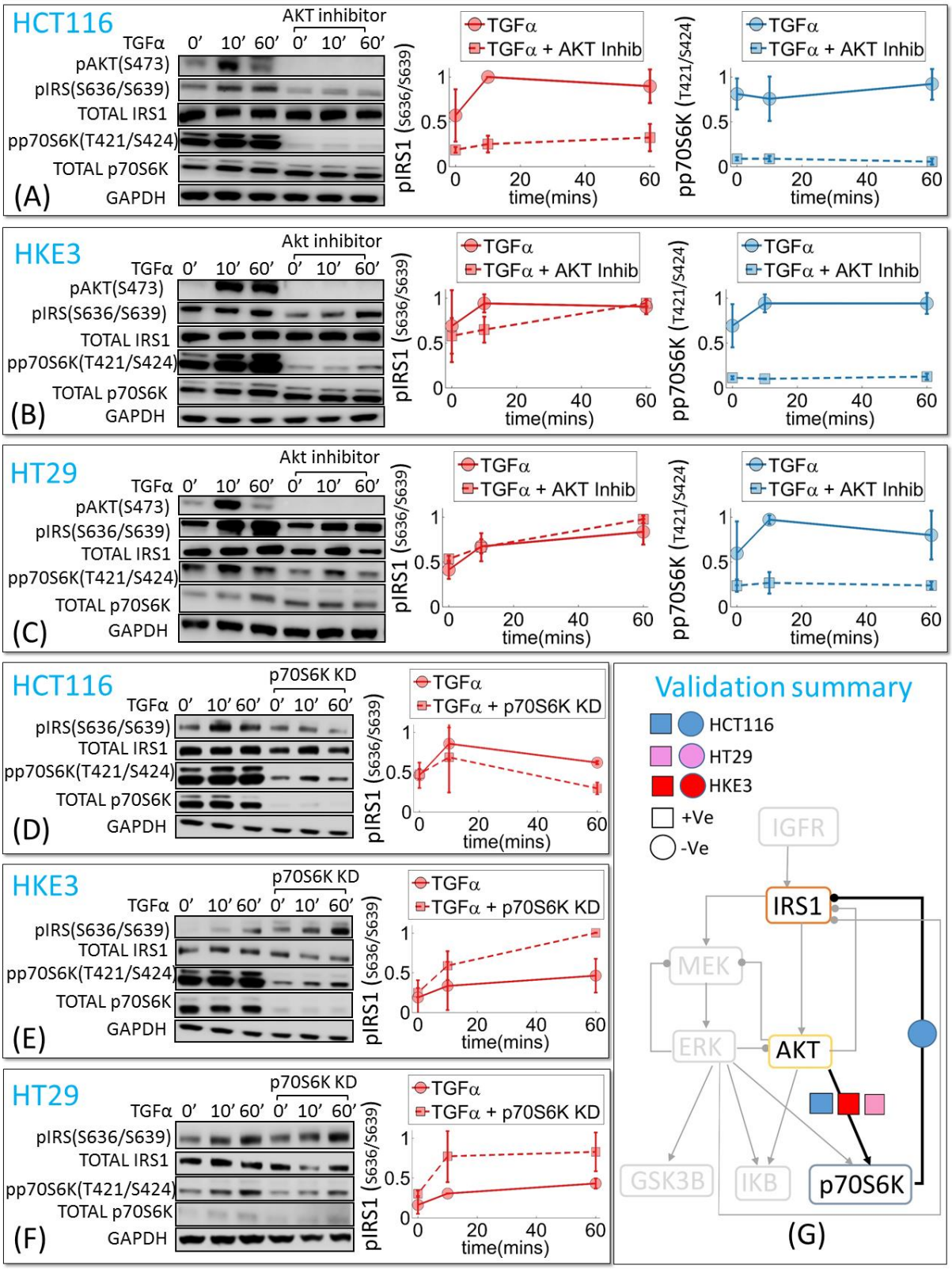
HCT116
AKT inhibitor
TGFα 0′ 10′ 60′ 0′ 10′ 60′
pAKT(S473)
pIRS(S636/S639)
TOTAL IRS1
pp70S6K(T421/S424)
TOTAL p70S6K
GAPDH
(A)
TGFα
TGFα + AKT Inhib
pIRS1 (S636/S639)
pp70S6K (T421/S424)
time(mins)
HKE3
Akt inhibitor
(B)
HT29
Akt inhibitor
(C)
HCT116
p70S6K KD
TGFα + p70S6K KD
(D)
HKE3
p70S6K KD
(E)
HT29
p70S6K KD
(F)
Validation summary
HCT116
HT29
HKE3
+Ve
-Ve
IGFR
IRS1
MEK
ERK
AKT
GSK3B
IKB
p70S6K
(G)

Figure 3

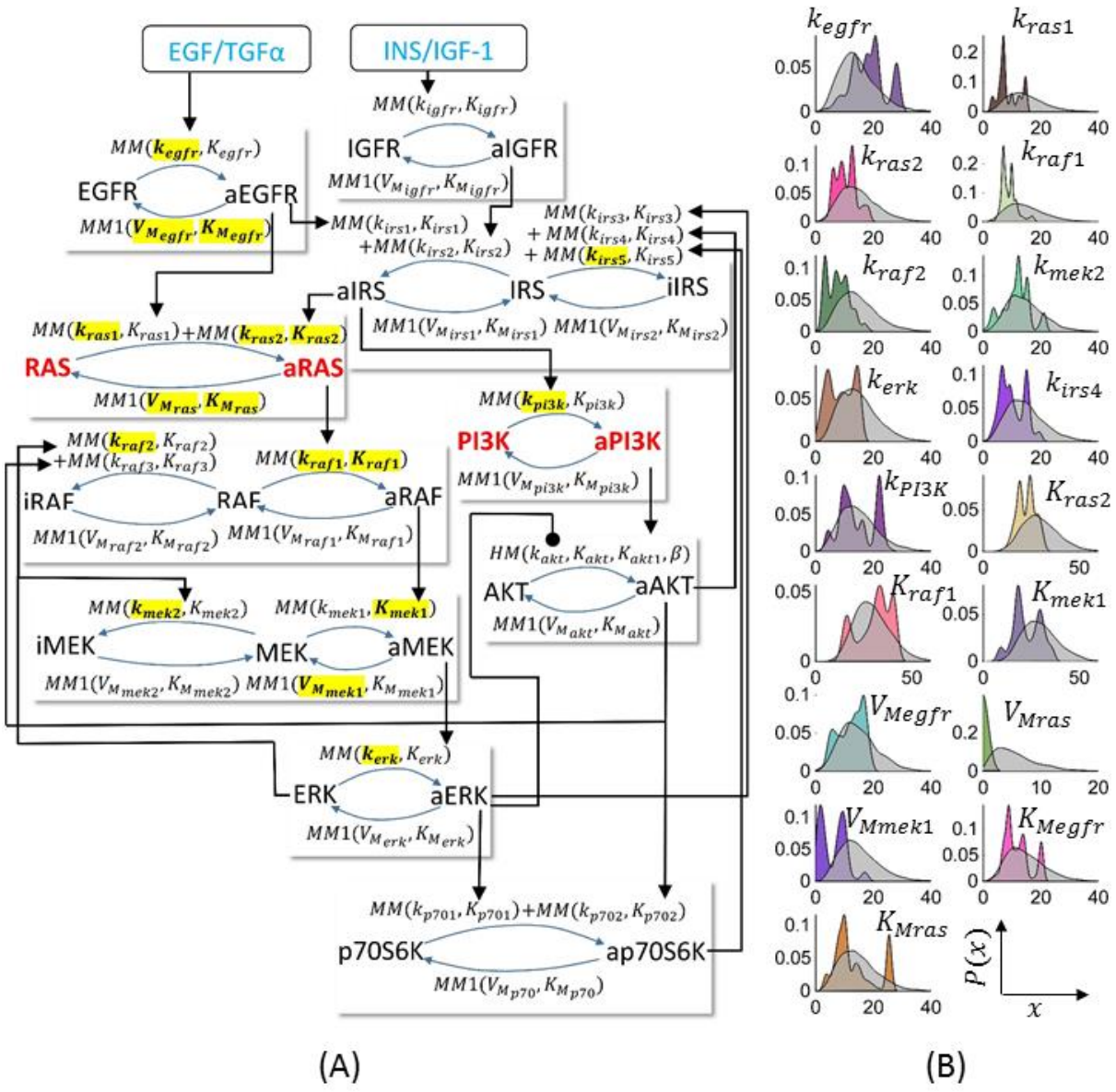
EGF/TGFα
INS/IGF-1
EGFR
aEGFR
IGFR
aIGFR
IRS
aIRS
iIRS
RAS
aRAS
PI3K
aPI3K
iRAF
RAF
aRAF
AKT
aAKT
iMEK
MEK
aMEK
ERK
aERK
p70S6K
ap70S6K
(A)
(B)

Figure 4

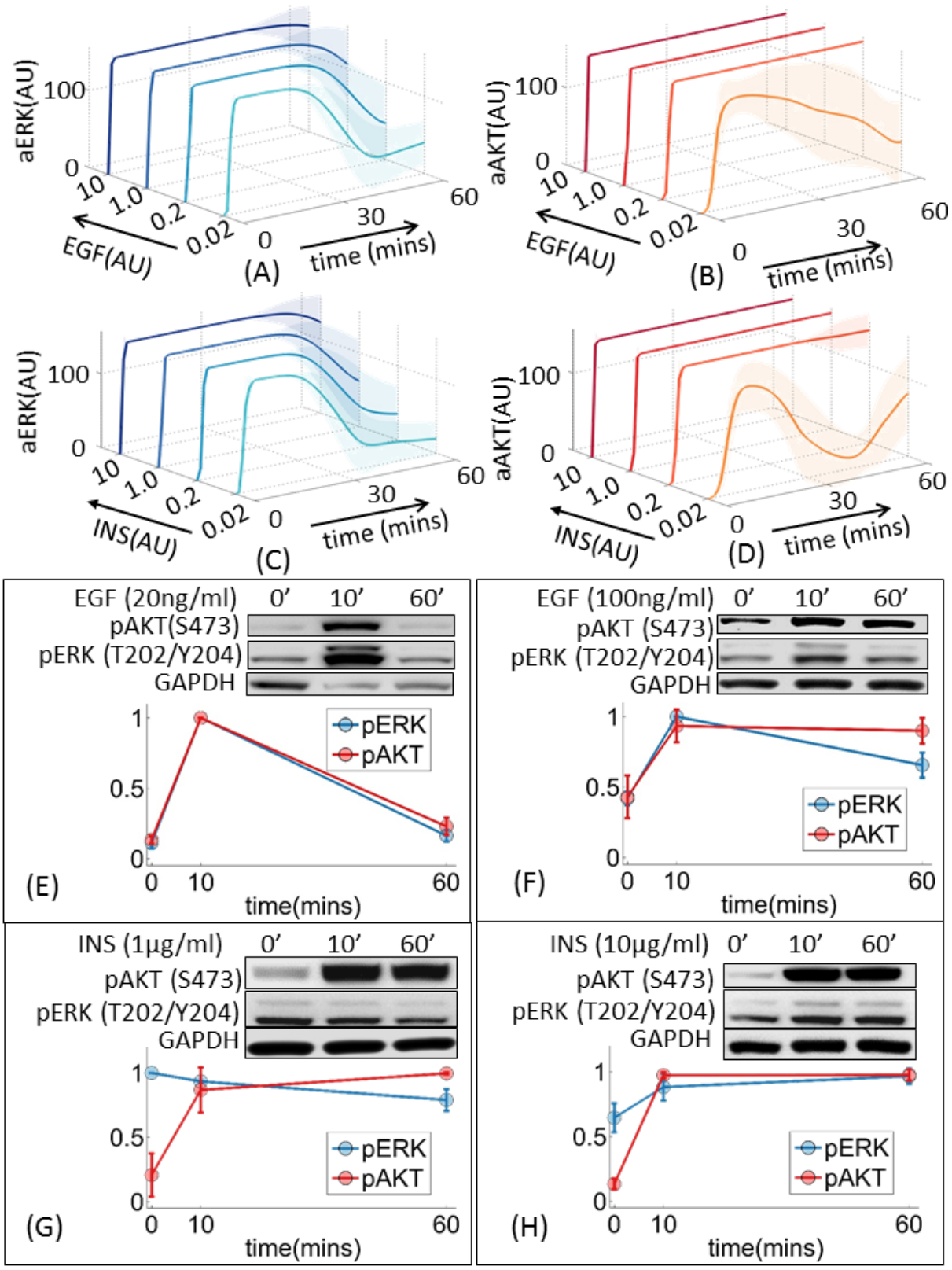
aERK(AU)
100
0
10
1.0
0.2
0.02
EGF(AU)
0
30
60
time (mins)
(A)
aAKT(AU)
(B)
INS(AU)
(C)
(D)
EGF (20ng/ml) 0' 10' 60'
pAKT(S473)
pERK (T202/Y204)
GAPDH
pERK
pAKT
1
0.5
0
time(mins)
(E)
EGF (100ng/ml)
(F)
INS (1µg/ml)
(G)
INS (10µg/ml)
(H)

Figure 5

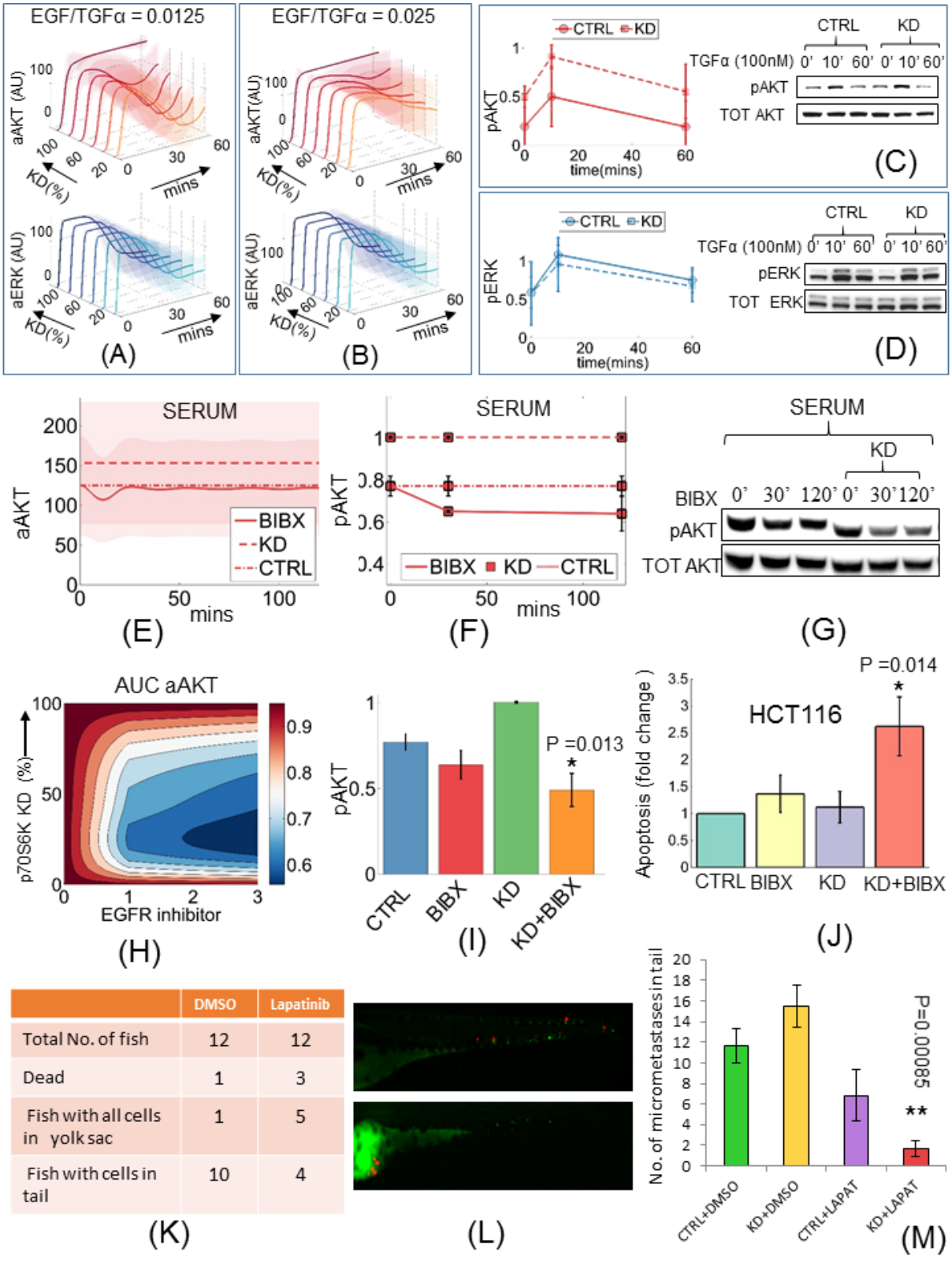
EGF/TGFα = 0.0125
EGF/TGFα = 0.025
aAKT (AU)
aERK (AU)
KD(%)
mins
(A)
(B)
CTRL KD
pAKT
time(mins)
TGFα (100nM) 0' 10' 60' 0' 10' 60'
pAKT
TOT AKT
(C)
pERK
TOT ERK
(D)
SERUM
BIBX
KD
CTRL
(E)
(F)
BIBX 0' 30' 120' 0' 30' 120'
(G)
AUC aAKT
p70S6K KD (%)
EGFR inhibitor
(H)
P =0.013
(I)
HCT116
Apoptosis (fold change)
P =0.014
CTRL BIBX KD KD+BIBX
(J)
DMSO
Lapatinib
Total No. of fish 12 12
Dead 1 3
Fish with all cells in yolk sac 1 5
Fish with cells in tail 10 4
(K)
(L)
No. of micrometastases in tail
P=0.00085
CTRL+DMSO
KD+DMSO
CTRL+LAPAT
KD+LAPAT
(M)

Figure 6