nature chemical biology

Identifying off-target effects and hidden phenotypes of drugs in human cells

Marnie L MacDonald[1], Jane Lamerdin[1], Stephen Owens[1], Brigitte H Keon[1], Graham K Bilter[1], Zhidi Shang[1], Zhengping Huang[1], Helen Yu[1], Jennifer Dias[1], Tomoe Minami[1], Stephen W Michnick[2] & John K Westwick[1]

We present a strategy for identifying off-target effects and hidden phenotypes of drugs by directly probing biochemical pathways that underlie therapeutic or toxic mechanisms in intact, living cells. High-content protein-fragment complementation assays (PCAs) were constructed with synthetic fragments of a mutant fluorescent protein ('Venus', EYFP or both), allowing us to measure spatial and temporal changes in protein complexes in response to drugs that activate or inhibit particular pathways. One hundred and seven different drugs from six therapeutic areas were screened against 49 different PCA reporters for ten cellular processes. This strategy reproduced known structure-function relationships and also predicted 'hidden,' potent antiproliferative activities for four drugs with novel mechanisms of action, including disruption of mitochondrial membrane potential. A simple algorithm identified a 25-assay panel that was highly predictive of antiproliferative activity, and the predictive power of this approach was confirmed with cross-validation tests. This study suggests a strategy for therapeutic discovery that identifies novel, unpredicted mechanisms of drug action and thereby enhances the productivity of drug-discovery research.

Drugs for various therapeutic indications frequently have 'hidden phenotypes' that result from unexpected or unintended activities, owing either to their binding to unknown targets or to unknown interactions between the intended drug target and other biochemical pathways. Such unexpected activities may either be harmful, leading to toxicity, or beneficial, suggesting new therapeutic indications. Observations of new and useful properties of drugs are usually made by serendipity, and the underlying mechanism by which a drug produces an effect is often not known. For example, thalidomide, which was originally used as an antiemetic agent, also has a broad spectrum of anticancer as well as teratogenic activities. The identification of hidden phenotypes has accelerated with the use of gene microarrays to identify groups of genes that are stimulated or repressed in response to specific conditions or treatments. Patterns of gene expression can provide a diagnostic fingerprint of drug activity, and in some cases they provide hypotheses about the cellular mechanisms of drug responses[1]. However, changes in mRNA expression may not correlate with the level or activity of any corresponding protein at a particular point in time, and the complexity of gene regulation can obscure the mechanisms of action of drugs within biochemical pathways[2,3]. In addition, activity of proteins is often dependent on cellular localization, information that is not captured by current genomic or proteomic analyses. These problems might be overcome if it were possible to probe specific events directly within biochemical pathways that underlie therapeutic or toxic mechanisms, and to do so in intact, living cells.

Here we present a strategy for identifying hidden phenotypes of drugs using high-content cellular analysis of changes in protein-protein interactions to probe individual steps in signal transduction pathways. For this purpose, we used PCAs based on enhanced and 'Venus' variants of yellow fluorescent protein (YFP)[4–8]. The principle of the PCA strategy is that cells simultaneously expressing two proteins fused to complementary fragments (F[1] and F[2]) of a reporter protein (enzyme) will produce a signal only if the fused proteins physically interact and bring the unfolded complementary fragments of the reporter protein into proximity, where they can fold together into an active enzyme. We demonstrate that this strategy not only reproduces known structure-function relationships for a series of drugs with known targets but also predicts phenotypes and mechanisms of action not previously described.

The strategy consists of six steps (**Fig. 1**). Briefly, we constructed PCAs for known protein complexes to act as reporters for activation or inhibition of diverse pathways (**Fig. 1**, step 1, described below). These assays have been validated in studies of diverse pathways, including insulin, growth factor–dependent, AKT kinase, erythropoietin, TGF-β and TNF signaling pathways[6,8–15]. Changes in protein complexes (interactions) may be effected by various biochemical events within a pathway, including post-translational modification, allosteric transition, protein degradation or *de novo* protein synthesis, protein conformational changes or protein translocation. Therefore, drugs that activate or inhibit the pathway in question can alter the amount of a particular complex, its subcellular location or both. Thus, we used a PCA based on intrinsically fluorescent proteins to capture quantitative, spatial and temporal changes in protein complexes. We measured changes in protein complexes in populations of cells grown

[1]Odyssey Thera, Inc. 4550 Norris Canyon Rd. Suite 140, San Ramon, California 94583, USA. [2]Département de Biochimie Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montreal, Quebec H3C 3J7, Canada. Correspondence should be addressed to S.W.M. (stephen.michnick@umontreal.ca) or J.K.W. (jwestwick@odysseythera.com).

Received 5 December 2005; accepted 4 April 2006; published online 7 May 2006; doi:10.1038/nchembio790

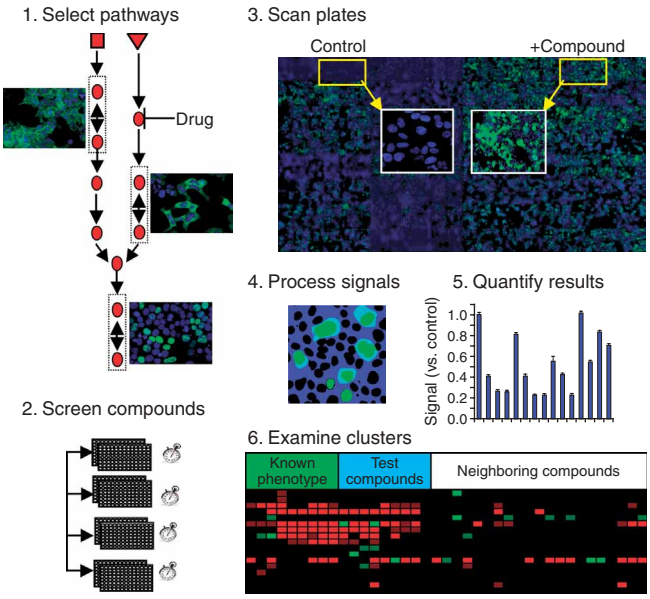

Figure 1 Strategy for pharmacological profiling of compounds with high-content PCAs. (1) Pathways of interest are selected and high-content PCA assays are created (pathway represented as red spheres connected by arrows). Assays measure dynamics of specific pathway activation or inhibition by quantifying changes in abundance or location of protein complexes coupled to that pathway that are elicited in response to activator (red square and triangle) or inhibitor drugs (⊥). Inset are images of three such assays that report on dynamic complexes coupled to the individual pathways (dotted-line boxes) localized to membrane, cytosol and nucleus. PCA signal is in green; nuclear (Hoechst) staining is in blue. (2) Cells expressing PCAs arrayed in 96-well plates are treated with compounds or vehicle controls, fixed after specified times and treated with cell compartment–specific counterstains. (3) Multiple images are captured from control and compound-treated wells. Pixel intensities from PCA signals are extracted from one or several cell compartments on the basis of colocalization with counterstain (4) and tabulated for individual compound treatments (5) (Methods and **Supplementary Methods**). Data for each compound versus PCA response at different times are represented as an array. Changes in signal intensity or location for compound versus vehicle control are represented by a color code, where green represents an increase and red a decrease in PCA signal versus control in units of coefficient of variation of each assay. Data are clustered by compounds and assays to identify on-pathway or off-pathway effects of compounds on specific pathways. The matrix also allows for the identification of test compounds that cluster with drugs of a known phenotype and could be expected to share the same phenotype.

on microtiter plates, using automated high-content microscopy and image analysis of the resulting fluorescence signals (**Fig. 1**, steps 2–4). To quantify drug effects, images of cells treated with vehicle or drugs for various lengths of time were analyzed by extracting the PCA signal from the regions of cells where the complex is formed. The signal was extracted using one of four different algorithms (steps 4 and 5), which were designed to quantify an increase, a decrease or a change in subcellular localization of the fluorescence signal (**Supplementary Methods** online). Finally, the quantitative effect of each drug on each assay was tabulated and subjected to hierarchical clustering of drugs and assays (step 6). In this way, individual clusters of drugs can be used to derive relationships between structure, function and assay activity. The matrix also allows for the identification of test compounds that cluster with drugs of a known phenotype and that could be expected to share the same phenotype, as opposed to neighboring compounds that share neither the same phenotype nor common assay responses. The predicted phenotypes of the test compounds can then be confirmed with functional assays in cells or model organisms of interest.

This strategy was implemented with PCAs for diverse, well-characterized protein interactions within cell-cycle, apoptosis, mitogenesis, ubiquitin-mediated proteolysis, G protein–coupled receptor (GPCR), chaperone, cytoskeletal, stress and inflammation, DNA-damage response and nuclear hormone receptor pathways (**Supplementary Table 1** online). We screened 107 different drugs representing six therapeutic areas (cancer, inflammation, cardiovascular disease, diabetes, neurological disorders and infectious disease) and cellular mechanisms (such as protein transport) at short and long time points in each of 49 different PCAs, resulting in 127 different measurements of each compound's activity. The resulting pharmacological profiles revealed both expected and unexpected effects of drugs, including four drugs—each of which is currently used for a completely different therapeutic indication—that clustered with drugs known to have antiproliferative activity. The previously hidden antiproliferative activity of these four drugs was confirmed in five different human tumor cell lines. A subset of only 25 assays was identified that would have predicted the antiproliferative phenotype with a high degree of likelihood (≥85%). Cross-validation tests based on leave-*n*-out analyses confirmed the predictability of the approach, with an average positive predictive value (PPV) of between 71% and 92%. Finally, individual assay signatures were found to correlate with specific cellular effects (for example, disruption of mitochondrial membrane potential) underlying the antiproliferative phenotype. This approach represents a novel process for therapeutic discovery in human cells.

RESULTS

Pharmacological profiling process

To create a set of PCA reporter assays, we selected proteins spanning a broad range of well-characterized cellular functions and drug target classes (**Supplementary Table 1**). 59 different proteins were used to construct 49 different assays, with each assay representing a pair of interacting proteins. Our goal was simply to place 'sensors' within pathways, avoiding perturbations of native signaling complexes. We therefore minimized the expression of exogenous proteins by engineering synthetic polypeptide fragments that generate an intense fluorescence signal upon complementation, even at near-endogenous levels of expression of the component proteins. This reconstituted reporter is 20 times brighter than YFP alone (**Supplementary Fig. 1** and **Supplementary Table 1** online).

HEK 293 cells that were transiently transfected with fusion constructs (**Supplementary Table 1**) or that stably expressed the constructs (see Methods) were plated in 96-well plates. Drugs were tested at concentrations (**Supplementary Table 2** online) that were based on cellular half-maximal inhibitory concentration (IC_{50}) values derived from the available literature for individual drugs, as described in Methods. Drug treatments were performed in duplicate wells with four scans per well, a procedure that captured images of a minimum of 800 cells per treatment. Well-to-well variations in transfection efficiencies were less than 5%. Fluorescence signals were acquired at a minimum of two wavelengths to detect the PCA signal and one or more counter-stains. An image from a portion of a scanned 96-well plate, in which the nuclei are counterstained with Hoechst (blue) and the PCA signal is in yellow-green, as shown in step 3 of **Figure 1**. Commonly observed drug effects included an increase, decrease or

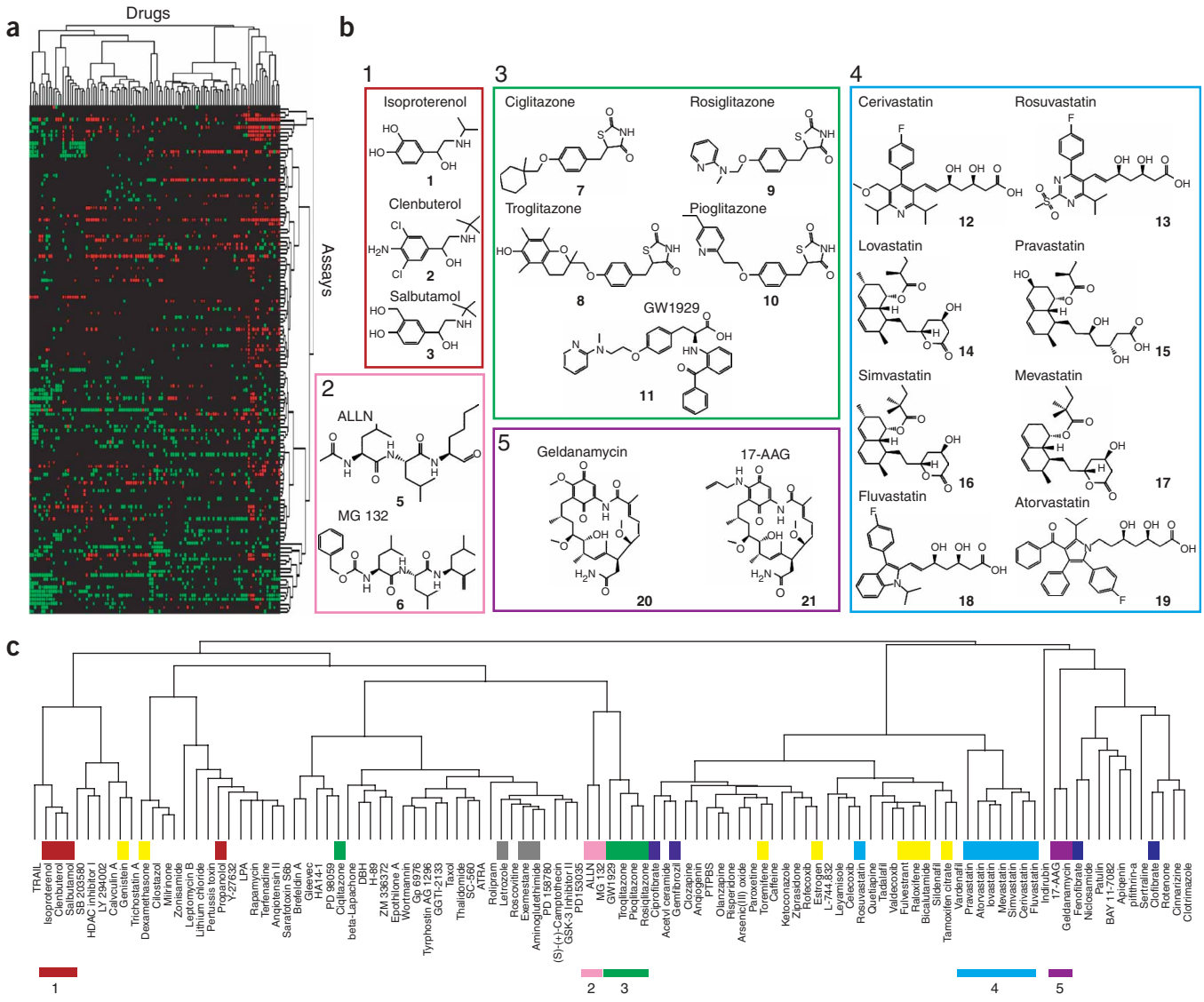

Figure 2 Drug activities by assay. (**a**) Matrix representation. Rows, 49 PCAs at individual time points (n = 127 total); columns, 107 different compounds; data have been hierarchically clustered in both dimensions (see Methods, **Table 1** and **Supplementary Tables 1** and **2**). Each data point is color coded to illustrate relative differences within an assay (Methods): green, increase relative to control; red, decrease; bright hue, difference of 2 × coefficient of variation; darker hue, difference of 1.5 × coefficient of variation. (**b**,**c**) Molecules of major drug clusters and expanded drug dendrogram (from **a**). Colored boxes below dendrogram indicate drug classes and major drug clusters, and structures of drugs in each cluster (**b**) are shown above the dendrogram (**c**). Major drug classes include adrenergic receptor agonists (1, red); proteasome inhibitors (2, pink); PPARγ glitazone/nonglitazone agonists (3, green); statins (4, light blue); and HSP90 inhibitors (5, purple). Also shown are steroid hormone receptor ligands (yellow), PPARα fibrate ligands (dark blue) and aromatase inhibitors (gray).

change in subcellular localization of the fluorescence signal at a particular time of treatment, and these changes were quantified (**Fig. 1**, steps 4 and 5) as described in **Supplementary Methods** online. Numerical results are shown in a color-coded matrix (**Figs. 1** and **2**), in which green denotes increases and red decreases in fluorescence intensity relative to a vehicle control, resulting from changes in either abundance or subcellular localization of the signal (see Methods for description of color intensities).

Drugs were tested at multiple time points, resulting in 127 assay–time point combinations for each drug and 13,589 data points overall (**Fig. 2** and **Supplementary Fig. 2** online). Unsupervised hierarchical clustering of both drugs and assays (columns and rows, respectively, in **Supplementary Fig. 2**) was applied to assess similarities between activity profiles (see Methods). Drugs that have similar activity across pathways will cluster together, and the underlying activity that drives the clustering can be seen (**Figs. 2** and **3**). If test compounds and known drugs form a cluster, they share common pathway activities that may be mechanistically related to the drugs' function and the associated phenotypes. Thus, the phenotypes of known drugs within a cluster provide clues as to the hidden phenotypes of the test compounds.

Chemically related drugs had similar activity profiles

Chemicals with similar structural features and molecular targets should have common effects on cellular pathways. Indeed, we

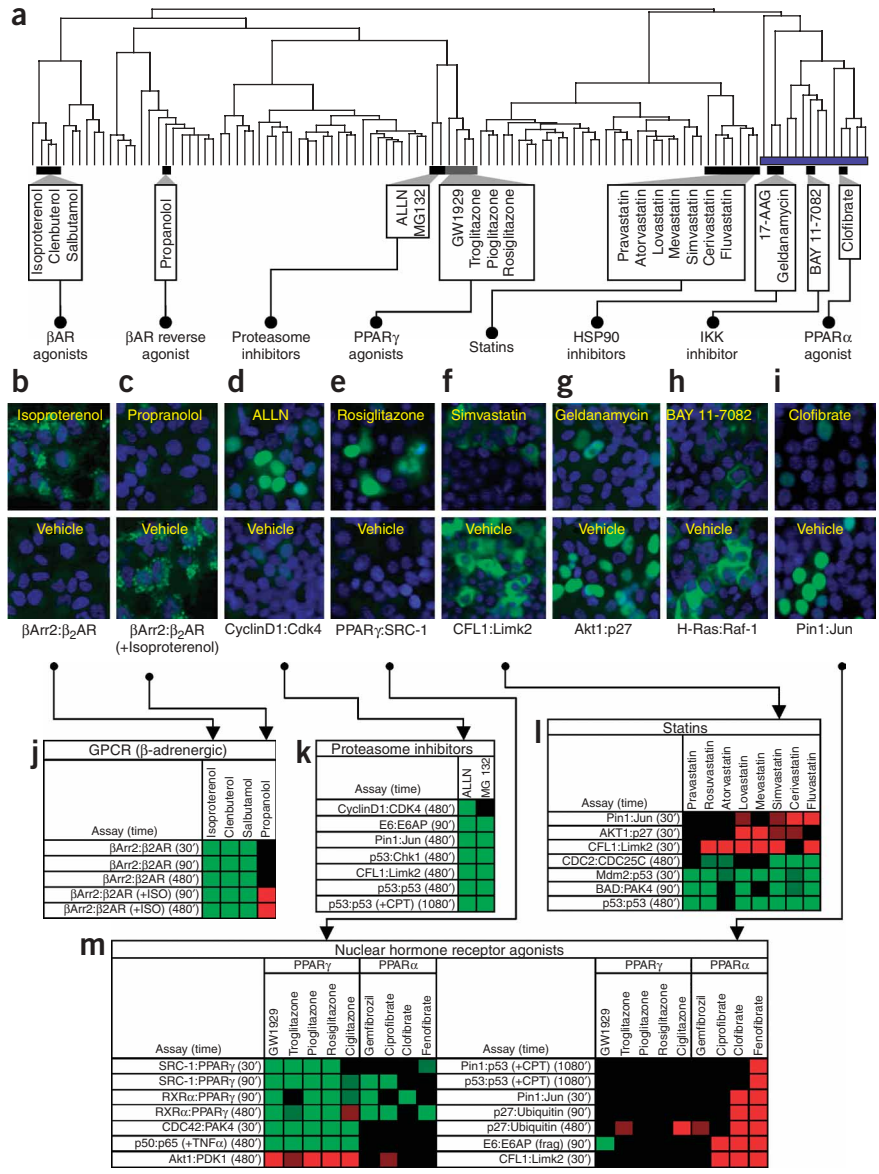

Figure 3 Clustering of compounds by assay response shows defined structure-function relationships. (**a**) Compound cluster dendrogram for 107 tested drugs (from **Fig. 2a**,**c**). Euclidean distance metrics were calculated for the log of the sample/control ratio. The blue bar corresponds to an inhibitory 'supercluster' discussed in **Figure 4**. (**b**–**i**) Individual structure classes and fluorescence micrographs showing examples of assays that respond to members of each structural class. (**j**–**m**) Matrices from **Figure 2a** of drug effects by compound structure–target class.

observed that structurally related drugs that share similar or identical cellular targets had very similar activities and pharmacological profiles. Major therapeutic target classes clustered together, revealing both 'on-target' and 'off-target' activities (**Figs. 2** and **3** and **Supplementary Fig. 2**).

We observed clustering of GPCR agonists, nuclear hormone receptor (peroxisome proliferator-activated receptor-γ, PPARγ) agonists, peptide proteasome inhibitors, statins and HSP90 inhibitors. One notable exception, the PPARα agonist fibrates, partitioned into two clusters, owing to distinct effects on different assays, as described below. In cases where the drug receptor or target was in the panel, the on-target activities of the drugs resulted in changes in known target protein complexes. For example, the β-adrenergic GPCR agonists, isoproterenol (compound **1**), clenbuterol (**2**) and salbutamol (**3**), all activated the β-adrenergic receptor and rapidly induced an interaction of the receptor with β-arrestin, as shown for isoproterenol in **Figure 3b**,**j**; this effect was blocked by pretreatment of cells with the inverse agonist propranolol (**4**) (**Fig. 3c**). The fluorescence micrograph of the isoproterenol-treated cells shows a punctate cytoplasmic pattern that is consistent with the binding of arrestin to the receptor and subsequent internalization via clathrin-coated pits[16], which is known to occur after ligand-induced phosphorylation of the receptor by GPCR kinases (GRKs)[17].

Assays reporting on the PPARs likewise responded in expected ways to known PPARγ agonists troglitazone (**8**), rosiglitazone (**9**), pioglitazone (**10**) and GW1929 (**11**). These compounds have a common phenylmethyl-2,4-thiazolidinedione core (except for GW1929). These agonists, which are known to promote the interaction of PPARγ with coactivators steroid receptor coactivator-1 (SRC-1) and retinoid X-receptor-α (RXRα)[18], induced target protein complexes (PPARγ:SRC-1 and PPARγ:RXRα) at short and long time points (**Fig. 3e**,**m**) and affected several other pathways as shown. In contrast, agonists of PPARα (gemfibrozil (**22**), ciprofibrate (**23**), fenofibrate (**24**) and clofibrate (**25**)) had different profiles (**Fig. 3m**). As PPARα was not present in this assay panel, the fibrates did not cluster together (**Fig. 2a**); nonetheless, they shared common off-target activities on several cell-cycle and apoptotic pathway assays, including Pin1:Jun (a nuclear complex containing the c-Jun transcription factor and the Pin1 prolyl isomerase; **Fig. 3i**). These results reveal both shared and distinct activities of the members of this drug class and also reveal, in the case of gemfibrozil, a certain degree of cross-reactivity with PPARγ. Ciprofibrate, clofibrate and fenofibrate seemed to be more selective than gemfibrozil, as shown in the left-hand matrix of **Figure 3m**.

The statins (the largest single class of compounds tested here) clustered well (**Figs. 2** and **3**) despite the absence of their known target, HMG–coenzyme A reductase, within the assay set. This clustering may result from an effect downstream of statins, as inhibition of HMG-CoA reductase also blocks the production of farnesyl pyrophosphate and geranyl-geranyl pyrophosphate. These pyrophosphates are required for isoprenylation of specific proteins, particularly in signal transduction pathways upstream of several of our reporter assays (**Fig. 3f**)[19].

On-target and off-target activities were also seen for the structurally related peptide proteasome inhibitors ALLN (**5**) and MG-132 (**6**), which had effects on known substrates of the proteasome (**Fig. 3d**,**k**)

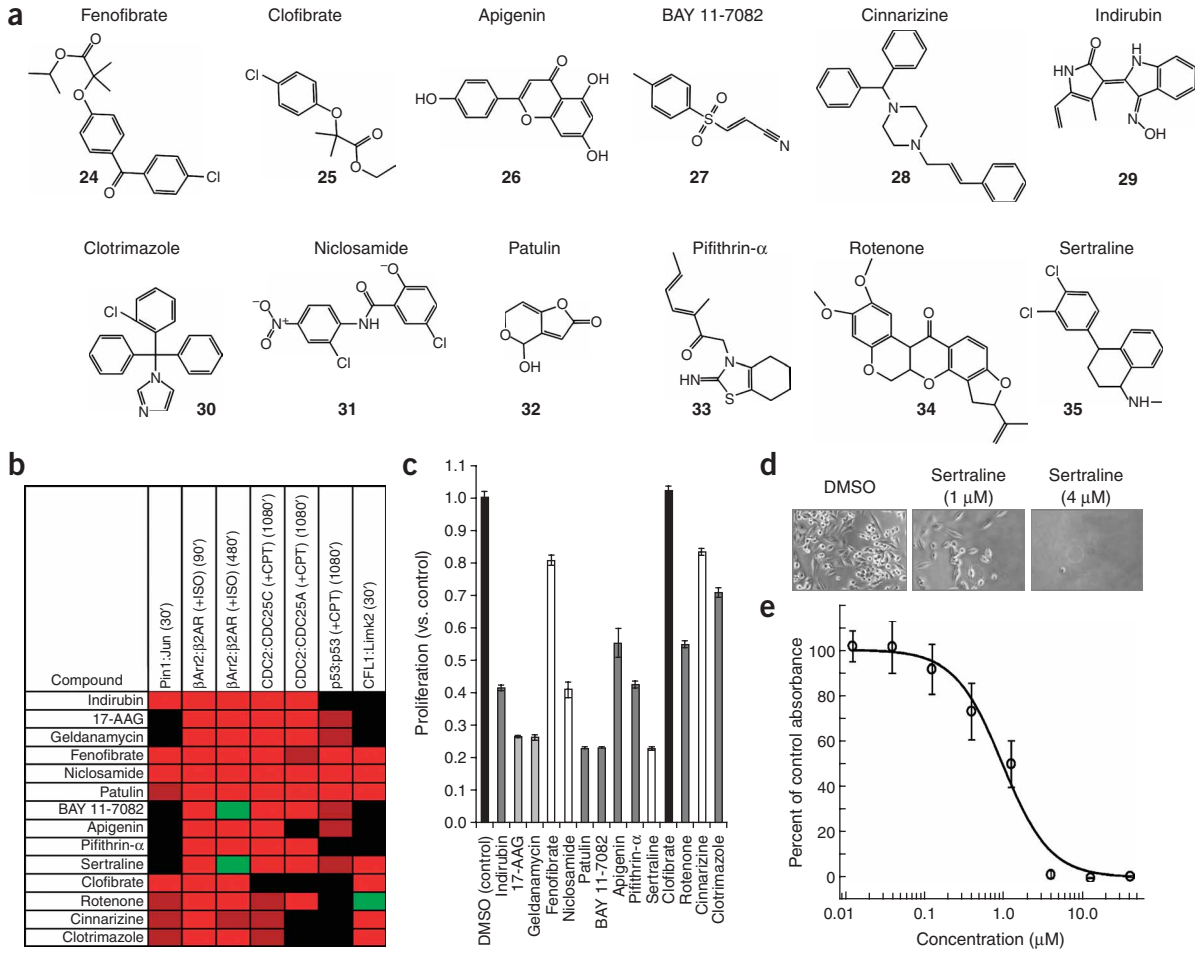

Figure 4 Supercluster from **Figure 3a** reveals compounds with potential hidden phenotypes. (**a**) Structures of compounds that clustered together on the basis of shared assay responses. (**b**) This drug cluster shares common assay responses and potentially similar cellular effects not obvious from structure or known indication, as shown in the matrix of drug effects, colored as in **Figure 2a**. (**c**) Compounds in the supercluster are antiproliferative in human prostate carcinoma cells (PC3). Results of MTT assays (**Supplementary Methods**) in PC3 cells treated with the indicated compounds. Black, control (DMSO) or non-antiproliferative; white, compounds with previously unknown antiproliferative activity; dark gray, compounds with known antiproliferative activity; light gray, HSP90 inhibitors. (**d**) Phase-contrast image of MiaPaCa cells grown in DMSO (vehicle) and two concentrations of sertraline. (**e**) Sertraline dose-response curve for proliferation of MiaPaCa cells as measured in an MTT assay (**Supplementary Methods**).

including p53, E6AP, c-Jun and cyclin D1 (refs. 20–23). Novel effects were also seen for these compounds; for example, complexes of cofilin with Lim kinase (CFL1:Limk2) accumulated in the presence of ALLN and MG132 (**Fig. 3k**), suggesting that a component of the complex—most probably Limk2—is regulated by the proteasome. In contrast, drugs that block protein chaperones would be expected to increase the turnover of their client proteins. Thus, we observed that the HSP90 inhibitor geldanamycin (**20**) and its semisynthetic analog 17-(allylamino)-17-demethoxygeldanamycin (17-AAG; **21**) affected numerous known HSP client protein complexes, resulting in a loss of signal (for example, see **Fig. 3g**).

Together, these results suggest that protein complexes respond dynamically to treatment with their intended drugs; both on-target and off-target activity of drugs can be studied by assessing protein complexes in human cells; and cellular selectivity of drugs can be assessed by studying complexes within different pathways. We demonstrated the utility of this strategy for assessing cellular selectivity of drugs across target classes and therapeutic areas. For example, we observed that the cyclin-dependent kinase (CDK) inhibitor indirubin-3′-monoxime (**29**) was strikingly nonselective, hitting numerous other target classes and pathways (**Supplementary Fig. 2**), consistent with published observations[24]. BAY 11-7082 (**27**), which was originally identified as an inhibitor of IKKγ, had effects on numerous assays, including a prototypical GTPase:kinase effector complex (H-Ras:Raf1; **Fig. 3h**). Nonselectivity in human cells is an important property of small molecules that may be assessed with this approach.

Hidden phenotypes and antiproliferative activities

The results discussed above suggest that similarities and differences between drugs, including on-target and off-target effects, can be identified. As off-target effects of drugs may be linked to hidden phenotypes, we investigated a 'supercluster' of drugs, marked with a blue bar in the right-hand side of the drug dendrogram (**Fig. 3a**) and shown in **Figure 4a**. These drugs had common effects on cytoskeletal, DNA-damage response, β-adrenergic and cell-cycle control pathways (**Fig. 4b**). We assessed whether the compounds shared the same therapeutic targets, purported mechanisms of action or structural features, but we found few similarities. Some of the compounds have

Table 1 IC_{50}s for the antiproliferative activity of drugs identified in the supercluster, as assessed in five different human tumor cell lines (described in Supplementary Methods).

		Cinnarizine	Fenofibrate	Niclosamide	Sertraline
	PC3	16.1	18.0	0.4	13.2
	A549	9.8	15.9	0.4	1.2
Antiproliferative	MiaPaCa	11.1	8.4	1.1	0.9
IC_{50} (μM)	LOVO	10.5	13.2	0.7	2.3
	U87MG	13.5	74.5	0.4	5.7
	Mean IC_{50} (μM)	10.2	30.0	0.6	4.7
Concentration (μM) used in HEK293 cells		15	30	1	10

known targets, whereas others act by unknown mechanisms. Some are in the current pharmacopoeia (fenofibrate (**24**), clofibrate (**25**), cinnarizine (**28**), clotrimazole (**30**) and sertraline (**35**)) and some are toxicants (patulin (**32**) and rotenone (**34**)). They had effects on Cdc2:Cdc25C, Cdc2:Cdc25A and p53:p53, in the presence of camptothecin (**40**), and on CFL1:Limk2 and Pin1:Jun (**Fig. 4b**). Many of the drugs in the cluster caused a reduction of Pin1:Jun complexes, including fenofibrate and clofibrate (**Figs. 3** and **4**), as discussed above.

Two common trends were observed: first, all of these compounds affected certain complexes in the cell cycle, DNA-damage response and/or cytoskeletal pathways, and these pathways are known to regulate cell growth; second, several drugs in the cluster are known to have antiproliferative activity. We hypothesized that the assay activity of these compounds that we observed is mechanistically related to their antiproliferative phenotypes. Thus, other drugs with similar assay profiles but an unknown phenotype might also have antiproliferative activity—a 'hidden phenotype' suggested by the activity profile in the cluster.

We tested this hypothesis by assessing the antiproliferative activity of the drugs in the supercluster in the human prostate carcinoma cell line (PC3). Geldanamycin, 17-AAG (**21**), apigenin (**26**), BAY 11-7082 (**27**), indirubin (**29**) and clotrimazole (**30**) all blocked the proliferation of PC3 cells, a known phenotype for these compounds, as did pifithrin-α (**Fig. 4c**)[24–28]. Consistent with our hypothesis, four other drugs in the supercluster were discovered to have previously unreported (hidden) antiproliferative activity. These included fenofibrate (Tricor), an antihyperlipidemic agent; niclosamide (**31**), an anthelminthic agent; cinnarizine (**28**), which is used to treat vertigo and other vestibular symptoms; and sertraline (**35**; Zoloft), which is used to treat depression. We tested each compound in a total of five human carcinoma cell lines derived from breast, colorectal, lung, prostate and pancreas (**Table 1**) and found IC_{50} values as low as 0.4 μM. Importantly, the average IC_{50} values for inhibition of proliferation in PC3 cells were very close to the doses used in pharmacological profiling in the HEK 293 cells (**Table 1**). Thus, effects on the PCAs were seen at doses that are relevant to the functional activity of these compounds.

Sertraline, a selective serotonin reuptake inhibitor, had potent and unexpected antiproliferative activity against a broad panel of tumor cell lines, with an average IC_{50} of 4.7 μM across the five different cell lines and an IC_{50} as low as 0.87 μM against a pancreatic carcinoma line (**Table 1** and **Fig. 4d,e**). Sertraline had strong inhibitory activity on multiple assays related to DNA-damage response pathways, as well as activity on the cytoskeletal signal-related probe CFL1:Limk2 (**Fig. 4b**). After the completion of these studies, a report confirmed the antiproliferative activity of sertraline and suggested a mechanism involving downregulation of a translationally controlled tumor protein[29]. In our studies, sertraline caused an early (30 min) and substantial (50%) reduction in the signaling complex Hsp90:eEF2K in addition to longer-term effects on HSP-related complexes including Hsp90:Cdc37 (a chaperone:cochaperone complex) (**Supplementary Fig. 2**). Decreased levels of active eEF2K are consistent with deregulation of the translational apparatus, and our results suggest a more general effect of sertraline on translation and protein homeostasis than that proposed in ref. 29. According to these results, the antiproliferative activity of sertraline seems to be an off-target effect that is not directly related to serotonin uptake.

Although fenofibrate was antiproliferative in PC3 cells, clofibrate—which was in the supercluster—had no effect on proliferation, and these two fibrates had distinct pharmacological profiles in our studies. Both agents are antihyperlipidemic, but fenofibrate had a marked effect on cell cycle and DNA-damage response assays, which may underlie its antiproliferative phenotype. Notably, fenofibrate (but not clofibrate) suppressed Cdc2:Cdc25C and p53:p53 in the presence of camptothesin (+CPT) (**Fig. 4b**).

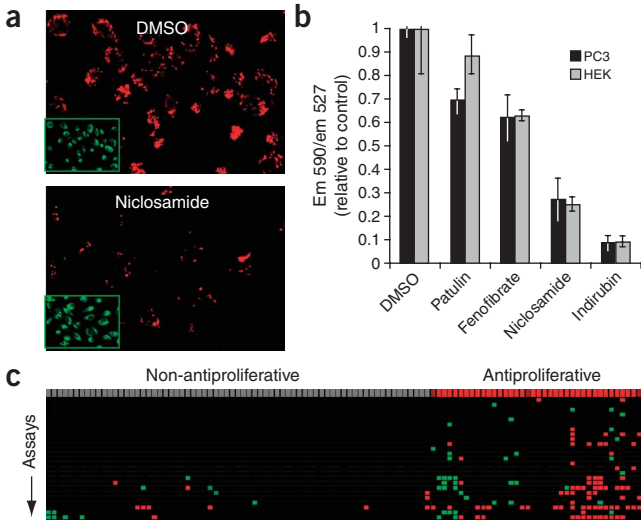

Figure 5 Identifying hidden phenotypes and underlying biochemical mechanisms. (**a**) Mitochondrial membrane potential was rapidly disrupted in cells treated with fenofibrate, patulin, niclosamide and indirubin; effect of niclosamide on mitochondria in PC3 cells is shown here. PC3 cells were treated with DMSO or drugs as indicated for 1 h and mitochondrial changes were assessed by staining with the cationic carbocyanine dye JC-1 (ref. 47 and **Supplementary Methods**). (**b**) Mitochondrial membrane potential in PC3 cells and HEK 293 cells. Results are presented as the ratio of red to green (em 590/em 527) ± s.e.m. ($n = 3$). Experiments were repeated at least three times in both cell lines at each time point with similar results. (**c**) Assay-response matrix (**Fig. 2a**) for a 25-assay subset, showing clear differentiation of antiproliferative compounds from non-antiproliferative compounds on the basis of assay responses. Assays were performed in HEK 293 cells (Methods) and proliferation was assessed in PC3 cells (**Supplementary Table 3** and **Supplementary Methods**). The 25 assays with the highest PPVs for antiproliferative activity were identified from the 127 assay-time combinations (**Fig. 2**) as described in Methods. Rows, top 25 assays, in the same order as in **Supplementary Table 4**; columns, 39 antiproliferative drugs and 68 non-antiproliferative drugs. See **Supplementary Figure 3** for matrix detail.

Table 2 Positive predictive values for antiproliferative activity, based on frequency of hits in the 25-assay subset.

	Antiproliferative		Non-antiproliferative		
Hits per compound (of 25 possible)	TPR	FNR	TNR	FPR	PPV
1	0.18	0.82	0.84	0.16	0.39
2	0.23	0.77	0.93	0.07	0.64
≥2	0.72	0.28	0.91	0.09	0.82
≥3	0.44	0.56	0.97	0.03	0.89
≥4	0.33	0.67	1.00	0.00	1.00

TPR, true positive rate for the antiproliferative phenotype; FNR, false negative rate; TNR, true negative rate; FPR, false positive rate; PPV, positive predictive value.

Several structurally distinct compounds in the supercluster (patulin, niclosamide, fenofibrate and indirubin) had notably similar activity profiles (**Fig. 4a**), suggesting similar cellular mechanisms of action. The penicillium mycotoxin patulin has previously been shown to inhibit cellular proliferation *in vitro* and tumor growth *in vivo*[30,31]. Niclosamide and fenofibrate have not previously been reported to block mammalian cell proliferation, yet both drugs had antiproliferative activity in five different human carcinoma cell lines, with average IC_{50} of 0.6 μM and 30 μM, respectively (**Table 1**). The anthelminthic activity of niclosamide has been linked to disruption of mitochondrial oxidative phosphorylation[32], and the toxicity of patulin has been linked to mitochondrial disruption[33]. Also, it has been shown that fibrates regulate fatty acid β-oxidation in peroxisomes and mitochondria[34,35] and that fenofibrate regulates mitochondrial membrane potential directly[36]. We investigated this phenotype in both HEK cells and PC3 cells and found that niclosamide, patulin and fenofibrate suppressed mitochondrial membrane potential within 1 h in both cell types (**Fig. 5a,b**), consistent with a direct action of these compounds on these processes. We observed that indirubin strongly inhibited mitochondrial membrane potential (90% inhibition within 1 h), a hidden phenotype that is unrelated to CDK, the purported target of indirubin[26]. Previous studies have suggested a transcriptional response of STAT3 and survivin as the mechanism underlying indirubin-induced apoptosis[37], but our results suggest a direct and previously unrecognized activity of this natural product on the mitochondria. Moreover, the concordant effects of these drugs on pathway activity, mitochondrial function and proliferation provides a direct link between the assay activity we observed and the functional consequences. The identification of a 'mitochondrial disruption signature' could facilitate the identification of other compounds sharing this activity.

Predictive assay subset

These results suggest that compounds with common underlying pathway activities cause similar phenotypes. Although the process of clustering allows the observation of overall similarities (and differences) from one compound to another, a particular phenotype may be related to effects on a specific pathway or, more likely, a particular combination of pathways. Therefore, we sought to determine whether a subset of assays would be capable of predicting the antiproliferative phenotype.

We first characterized all 107 drugs for their ability to inhibit the proliferation of PC3 cells and identified 39 drugs that were antiproliferative and 68 drugs that were not antiproliferative (**Supplementary Table 3** online). We then used this phenotype to identify the most predictive of the 127 assays by calculating a PPV (Methods) for each assay[38,39]. For any one assay, a high PPV means a high likelihood of detection of a true positive (antiproliferative) drug versus a false positive (non-antiproliferative) drug. We calculated the PPV for each assay and time point and found that a panel of only 25 assays could have predicted the antiproliferative phenotype, with a combined PPV of 85% for the 25-assay subset (**Supplementary Table 4** online), and that antiproliferative drugs could clearly be distinguished from non-antiproliferative drugs with these 25 assays (**Fig. 5c** and **Supplementary Fig. 3** online). The 25-assay subset was comprised primarily of proteins known to be involved in cell cycle, apoptosis, cytoskeletal and DNA-damage response pathways. Although many compounds in the supercluster had secondary activity on β-arrestin in the presence of isoproterenol (**Fig. 4b**), this activity seemed to represent a common off-target effect that was not correlated with cell proliferation, as the β-arrestin assay was not predictive of the antiproliferative phenotype (PPV = 50% at 90 min).

Because proliferation is a phenotype that involves multiple cellular mechanisms[40], a drug that hits a combination of proliferation-dependent pathways should have a higher likelihood of being antiproliferative. We can assess the likelihood that a drug would be antiproliferative on the basis of a certain hit frequency, using the equation described in the Methods to evaluate drugs that hit one or more assays in the 25-assay subset (**Table 2**).

Accordingly, we observed that the PPV increased as the frequency of hits increased. Drugs that hit two or more of these 25 assays had an 82% likelihood of being antiproliferative, and drugs that hit three or more of these assays had an 89% likelihood of being antiproliferative (**Table 2**). Finally, 13 drugs hit four or more assays in the panel, and all 13 of these were antiproliferative, including indirubin, geldanamycin, fenofibrate, niclosamide, patulin, BAY 11-7082, apigenin, pifithrin-α and sertraline. Different antiproliferative drugs had different signatures on the assay panel. For example, LY 294002 (**36**), trichostatin A (**37**), calyculin A (**38**) and HDAC inhibitor I (**39**) had positive effects on several cell-cycle pathways in a manner opposite to the drugs in the supercluster (**Supplementary Fig. 3**) and had no effect on mitochondrial function, suggesting a completely different mechanism of action.

In the analyses described above, all 107 drugs were used to estimate the predictive power of the assays (**Table 2** and **Supplementary Table 4**). To perform leave-*n*-out cross-validation tests, drugs were randomly separated into training sets and validation sets, and the training set(s) of drugs were used to identify the most predictive assays as described in the **Supplementary Methods**. That assay panel was then tested against the corresponding validation set(s) of drugs. The leave-*n*-out analysis was performed for $n = 2, 3, \ldots, 10$, where n is the number of training sets, and the results of all the analyses were combined to calculate the predictive power of the assays for the antiproliferative phenotype.

The detailed results are shown in **Supplementary Table 4** and are summarized here. The PPV obtained from the leave-*n*-out analysis was 71.3% for at least 1 of 25 possible hits. Therefore, any drug that hit one or more assays in a 25-assay panel had a 71% likelihood of being antiproliferative (as described in **Supplementary Methods**, we call this the 'unconstrained' case). However, as shown in **Table 2**, the likelihood of detecting the antiproliferative phenotype increased as the frequency of hits increased (we call this the 'constrained' case). The leave-*n*-out analysis showed that the average PPV in the constrained case was 91.8% (**Supplementary Table 4**). The results of these cross-validation tests support the conclusion that the assays are predictive of the antiproliferative phenotype.

These assays can be used to discover additional compounds with these phenotypes and to identify specific biochemical mechanisms

underlying the phenotypes. The assays used here were not originally chosen to probe pathways linked to any particular phenotype. It is likely, therefore, that even higher predictive power could be obtained by designing an assay panel focused on the underlying mechanisms of action of antiproliferative compounds. Finally, a similar overall strategy could be used to identify many other hidden phenotypes linked to function, as discussed below.

DISCUSSION

Numerous drugs routinely used as human therapeutics act, at least in part, by unknown mechanisms or have hidden phenotypes. We have observed here and in previous studies that unanticipated drug effects in a model cell line can predict important phenotypes in cells of different lineages, even if the functional phenotype is not observed in the model cell line used[6,11,12]. By comparing the profiles of these drugs within and between target classes, an understanding of the proteins and pathways contributing to drug activity can be obtained and hidden phenotypes may be predicted. These hidden phenotypes can point to unforeseen and potentially useful applications of drugs and to the formulation of new hypotheses about drug actions. In addition, most, if not all, therapeutic agents, and in particular low–molecular weight chemicals, bind macromolecules other than their intended targets. Current strategies do not enable the identification of these off-target events. We observed these effects for numerous drug classes, including the fibrates, glitazones and statins (**Fig. 3l**,**m**), and for nonselective inhibitors such as indirubin.

The results of this study suggest advances toward generalization of this approach. First, we were able to predict a very general phenotype (antiproliferative activity) by using a standard cell line to perform the screens (HEK 293), yet the resulting phenotypes were observed in human tumor cells. Further, the approach is not limited to the identification of compounds likely to generate a gross effect, as we showed that particular biochemical mechanisms, such as mitochondrial membrane potential disruption, are consistent with specific pathway activities and were manifested both in HEK cells and in PC3 cells (**Fig. 5b**). Many of the well-characterized pathways we probed here are undoubtedly conserved in many cell types, such that the screens could be performed in an HEK cell line to identify the underlying pathway activity, as we did here, and the predicted phenotype observed in another cell line that manifests the phenotype. By contrast, for tissue-specific pathways, the screens themselves could be performed in more specialized primary or immortalized cells. We and others have successfully performed PCAs in a variety of mammalian, bacterial, yeast and insect cell lines and in primary cells and whole organisms[8,11–14,41–45].

Many of the proteins studied here are known to be involved in cell proliferation–related pathways, so the links we observed between assay activity and that phenotype are perhaps not surprising. Nevertheless, the universality of the method suggests that assays can be developed for many other interactions in pathways linked to a wide range of disease processes, targets, metabolic and signaling pathways, and therapeutic indications. Such assays could include proteins that are expressed only in a particular differentiated cell type.

Hidden phenotypes related to toxicity are responsible for the high failure rates of drug candidates, and improved methods for identifying these unintended effects are needed. By carrying out the pharmacological profiling process with toxicants and failed compounds, of both known and unknown mechanisms of action, predictive assays could be established for these undesirable activities. The principles described here should provide an efficient means to flag compounds for exclusion from further studies or to redirect development to other therapeutic areas (a 'fail-fast' strategy). In addition, understanding the specific biochemical nature of the off-target activity of a lead compound would enable optimization of a chemical structure to enhance desirable attributes and avoid undesirable attributes. We observed clear differences in activity between closely related compounds in a structural class, suggesting that minor chemical modifications result in differences in compound activity in living cells. These strategies may help to clarify our understanding of drug action and enhance the productivity of drug-discovery research.

METHODS

Assay selection, constructions and transfections. The 49 different protein-fragment complementation assays are shown in **Supplementary Table 1**. The use of a rapidly maturing, intensely fluorescent mutant of YFP known as Venus[7] allowed for low levels of expression in human cells (**Supplementary Fig. 1**). PCA fragments were obtained as follows: first, fragments coding for YFP[1]and YFP[2] (corresponding to amino acid residues 1–158 and 159–239 of YFP, respectively) were generated by oligonucleotide synthesis (Blue Heron Biotechnology), and then PCR mutagenesis was used to generate the mutant fragments IFP[1] and IFP[2]. The IFP[1] fragment corresponds to YFP[1] F46L F64L M153T and the IFP[2] fragment corresponds to YFP[2] V163A S175G. Fusion constructs were generated as previously described[8].

HEK 293 cells were maintained as recommended by ATCC. For p50:p65, βArr2:β2AR and Akt1:PDK1, stable cell lines were generated as described previously[8] and seeded 24 h before drug treatment. For all other PCAs, HEK 293 cells were seeded into 96-well poly-D-lysine–coated plates (Greiner) 24 h before transfection and then cotransfected with between 10 and 100 ng of the vector pairs, using Fugene 6 (Roche) according to the manufacturer's protocol. The lowest possible DNA concentration generating an adequate signal was chosen, ensuring correct subcellular localization of the resulting complexes. Beginning at either 24 or 48 h after transfection, cells were screened against drugs as described below.

Drug screening. We screened 107 different drugs against 49 distinct assays in duplicate at one or more time points. Drug concentrations were chosen to approximate three times the published cellular IC_{50} values (**Supplementary Table 2**) and were tested to ensure lack of overt toxicity in HEK 293 cells. Cells expressing the PCA pairs were incubated in medium containing drugs for 30 min, 90 min and 8 h, or, in the case of DNA-damage response pathways, for 18 h. For certain pathways, cells were also stimulated with agonist (CPT, TNFα or isoproterenol) as indicated in **Supplementary Table 1**. For CPT, cells were first treated with drugs for 2 h and then CPT was added for 16 h. For TNFα and isoproterenol, cells were first treated with drugs and the agonist was added during the last 30 and 60 min of incubation, respectively. After the drug and agonist treatments, the cells were simultaneously stained with Hoechst 33342 (Molecular Probes) and 2% (v/v) formaldehyde (Ted Pella). For Akt1:PDK1, cells were also stained with 15 μg ml^{-1} Texas Red–conjugated wheat-germ agglutinin (Molecular Probes). Cells were then rinsed with Hank balanced salt solution (Invitrogen) and maintained in the same buffer during image acquisition. Four scans were acquired for each well, with each scan representing 150–300 cells. A constant exposure time for each wavelength was used to acquire all images for a given assay. Fluorescence image analysis was performed as described in **Supplementary Methods**.

Data visualization and clustering. In the drug matrix (**Fig. 2**) each column represents one drug and each row represents one assay at a particular time point in the absence or presence of an agonist treatment. Each data point was formed by taking the log of the ratio of the sample signal to the control signal. Every row and column carries equal weight. The Ward hierarchical clustering algorithm[46] and euclidean distance metrics (http://www.r-project.org/) were used for clustering the results. For display purposes, each data point in the matrix is color coded to illustrate relative differences within an assay, expressed as a multiple of the coefficient of variation of the controls in each assay. An increase relative to the control (vehicle-only treatment) is displayed as green and a decrease relative to the control is displayed as red. Each color is further divided into two levels: level 1 (2 × coefficient of variation) and level 2 (1.5 ×

coefficient of variation). Level 1 is displayed as the brighter hue and level 2 as the darker hue.

Identification of a predictive assay panel. Each assay at each time point ($n = 127$) was evaluated for its sensitivity and specificity with respect to phenotype (antiproliferative versus non-antiproliferative) using standard screening methodologies[38,39]. First, any level 1 hit (as defined above) was assigned a score of 1, regardless of sign (positive or negative). Weaker (level 2) hits were ignored. We then calculated the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) rates for each assay. 'Positive' refers to the antiproliferative phenotype and 'negative' refers to the non-antiproliferative phenotype, where cell proliferation was assessed by a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (**Supplementary Methods**). The PPV was calculated as follows:

$$\mathrm{PPV} = \mathrm{TP}_i/(\mathrm{TP}_i + \mathrm{FP}_i)$$

where $\mathrm{TP}_i + \mathrm{FP}_i \neq 0$. The assays were then sorted by PPV in descending order to identify the most predictive assays (**Supplementary Table 4**). In the above analysis, the frequency of hits that any drug made was not considered when assessing the predictive nature of the assays (unconstrained case). Considering that a drug that hits more than one proliferation-related pathway may have a greater likelihood of being antiproliferative, we also calculated the screening results on the basis of hit frequency by classifying antiproliferative and non-antiproliferative drugs according to their activity on either ≥ 1, ≥ 2, ≥ 3 or ≥ 4 assays of the 25-assay panel (constrained case) (**Table 2**). Cross-validation tests based on leave-*n*-out analyses were performed as described in **Supplementary Methods**.

Note: Supplementary information is available on the Nature Chemical Biology website.

ACKNOWLEDGMENTS

The authors thank E. Smith, M. West and the molecular biology staff at Odyssey Thera for excellent technical support. S.W.M. is the Canada Research Chair in Integrative Genomics. This manuscript is dedicated to the memory of Anthony V. Carrano.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturechemicalbiology
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Stoughton, R.B. & Friend, S.H. How molecular profiling could revolutionize drug discovery. *Nat. Rev. Drug Discov.* **4**, 345–350 (2005).
2. Tian, Q. *et al.* Integrated genomic and proteomic analyses of gene expression in Mammalian cells. *Mol. Cell. Proteomics* **3**, 960–969 (2004).
3. Miklos, G.L. & Maleszka, R. Microarray reality checks in the context of a complex disease. *Nat. Biotechnol.* **22**, 615–621 (2004).
4. Ghosh, I., Hamilton, A.D. & Regan, L. Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *J. Am. Chem. Soc.* **122**, 5658–5659 (2000).
5. Michnick, S.W., Remy, I., Campbell-Valois, F.X., Vallee-Belisle, A. & Pelletier, J.N. Detection of protein-protein interactions by protein fragment complementation strategies. *Methods Enzymol.* **328**, 208–230 (2000).
6. Remy, I. & Michnick, S.W. Visualization of biochemical networks in living cells. *Proc. Natl. Acad. Sci. USA* **98**, 7678–7683 (2001).
7. Nagai, T. *et al.* A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87–90 (2002).
8. Yu, H. *et al.* Measuring drug action in the cellular context using protein-fragment complementation assays. *Assay Drug Dev. Technol.* **1**, 811–822 (2003).
9. Remy, I. & Michnick, S.W. Mapping biochemical networks with protein-fragment complementation assays. *Methods Mol. Biol.* **261**, 411–426 (2004).
10. Remy, I. & Michnick, S.W. A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways. *Methods* **32**, 381–388 (2004).
11. Remy, I. & Michnick, S.W. Regulation of apoptosis by the Ft1 protein, a new modulator of protein kinase B/Akt. *Mol. Cell. Biol.* **24**, 1493–1504 (2004).
12. Remy, I., Montmarquette, A. & Michnick, S.W. PKB/Akt modulates TGF-beta signalling through a direct interaction with Smad3. *Nat. Cell Biol.* **6**, 358–365 (2004).
13. Nyfeler, B., Michnick, S.W. & Hauri, H.P. Capturing protein interactions in the secretory pathway of living cells. *Proc. Natl. Acad. Sci. USA* **102**, 6350–6355 (2005).
14. Hu, C.D., Chinenov, Y. & Kerppola, T.K. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell* **9**, 789–798 (2002).
15. Fang, D. & Kerppola, T.K. Ubiquitin-mediated fluorescence complementation reveals that Jun ubiquitinated by Itch/AIP4 is localized to lysosomes. *Proc. Natl. Acad. Sci. USA* **101**, 14782–14787 (2004).
16. Zhang, J., Ferguson, S.S., Barak, L.S., Menard, L. & Caron, M.G. Dynamin and beta-arrestin reveal distinct mechanisms for G protein-coupled receptor internalization. *J. Biol. Chem.* **271**, 18302–18305 (1996).
17. Lin, F.T. *et al.* Phosphorylation of beta-arrestin2 regulates its function in internalization of beta(2)-adrenergic receptors. *Biochemistry* **41**, 10692–10699 (2002).
18. Nolte, R.T. *et al.* Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature* **395**, 137–143 (1998).
19. Winter-Vann, A.M. & Casey, P.J. Post-prenylation-processing enzymes as new targets in oncogenesis. *Nat. Rev. Cancer* **5**, 405–412 (2005).
20. Maki, C.G., Huibregtse, J.M. & Howley, P.M. In vivo ubiquitination and proteasome-mediated degradation of p53(1). *Cancer Res.* **56**, 2649–2654 (1996).
21. Treier, M., Staszewski, L.M. & Bohmann, D. Ubiquitin-dependent c-Jun degradation in vivo is mediated by the delta domain. *Cell* **78**, 787–798 (1994).
22. Nefsky, B. & Beach, D. Pub1 acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25. *EMBO J.* **15**, 1301–1312 (1996).
23. Murray, A. Cyclin ubiquitination: the destructive end of mitosis. *Cell* **81**, 149–152 (1995).
24. Bain, J., McLauchlan, H., Elliott, M. & Cohen, P. The specificities of protein kinase inhibitors: an update. *Biochem. J.* **371**, 199–204 (2003).
25. Fotsis, T. *et al.* Flavonoids, dietary-derived inhibitors of cell proliferation and in vitro angiogenesis. *Cancer Res.* **57**, 2916–2921 (1997).
26. Hoessel, R. *et al.* Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. *Nat. Cell Biol.* **1**, 60–67 (1999).
27. Benzaquen, L.R., Brugnara, C., Byers, H.R., Gatton-Celli, S. & Halperin, J.A. Clotrimazole inhibits cell proliferation in vitro and in vivo. *Nat. Med.* **1**, 534–540 (1995).
28. Neckers, L., Schulte, T.W. & Mimnaugh, E. Geldanamycin as a potential anti-cancer agent: its molecular target and biochemical activity. *Invest. New Drugs* **17**, 361–373 (1999).
29. Tuynder, M. *et al.* Translationally controlled tumor protein is a target of tumor reversion. *Proc. Natl. Acad. Sci. USA* **101**, 15364–15369 (2004).
30. Keblys, M. *et al.* The effects of the Penicillium mycotoxins citrinin, cyclopiazonic acid, ochratoxin A, patulin, penicillic acid, and roquefortine C on in vitro proliferation of porcine lymphocytes. *Mycopathologia* **158**, 317–324 (2004).
31. Seigle-Murandi, F., Steiman, R., Krivobok, S., Beriel, H. & Benoit-Guyod, J.L. Anti-tumor activity of patulin and structural analogs. *Pharmazie* **47**, 288–291 (1992).
32. Weinbach, E.C. & Garbus, J. Mechanism of action of reagents that uncouple oxidative phosphorylation. *Nature* **221**, 1016–1018 (1969).
33. Burghardt, R.C. *et al.* Patulin-induced cellular toxicity: a vital fluorescence study. *Toxicol. Appl. Pharmacol.* **112**, 235–244 (1992).
34. Minnich, A., Tian, N., Byan, L. & Bilder, G. A potent PPARα agonist stimulates mitochondrial fatty acid beta-oxidation in liver and skeletal muscle. *Am. J. Physiol. Endocrinol. Metab.* **280**, E270–E279 (2001).
35. Froyland, L. *et al.* Mitochondrion is the principal target for nutritional and pharmacological control of triglyceride metabolism. *J. Lipid Res.* **38**, 1851–1858 (1997).
36. Zhou, S. & Wallace, K.B. The effect of peroxisome proliferators on mitochondrial bioenergetics. *Toxicol. Sci.* **48**, 82–89 (1999).
37. Nam, S. *et al.* Indirubin derivatives inhibit Stat3 signaling and induce apoptosis in human cancer cells. *Proc. Natl. Acad. Sci. USA* **102**, 5998–6003 (2005).
38. Wald, N. & Cuckle, H. Reporting the assessment of screening and diagnostic tests. *Br. J. Obstet. Gynaecol.* **96**, 389–396 (1989).
39. Parsonnet, J. & Axon, A.T. Principles of screening and surveillance. *Am. J. Gastroenterol.* **91**, 847–849 (1996).
40. Hanahan, D. & Weinberg, R.A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
41. Galarneau, A., Primeau, M., Trudeau, L.E. & Michnick, S.W. β-Lactamase protein fragment complementation assays as *in vivo* and *in vitro* sensors of protein protein interactions. *Nat. Biotechnol.* **20**, 619–622 (2002).
42. Leveson-Gower, D.B., Michnick, S.W. & Ling, V. Detection of TAP family dimerizations by an in vivo assay in mammalian cells. *Biochemistry* **43**, 14257–14264 (2004).
43. Pelletier, J.N., Campbell-Valois, F. & Michnick, S.W. Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl. Acad. Sci. USA* **95**, 12141–12146 (1998).
44. Subramaniam, R., Desveaux, D., Spickler, C., Michnick, S.W. & Brisson, N. Direct visualization of protein interactions in plant cells. *Nat. Biotechnol.* **19**, 769–772 (2001).
45. Benton, R., Sachse, S., Michnick, S.W. & Vosshall, L.B. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biol.* **4**, e20 (2006).
46. Ward, J. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58**, 236–244 (1963).
47. Reers, M., Smith, T.W. & Chen, L.B. J-aggregate formation of a carbocyanine as a quantitative fluorescent indicator of membrane potential. *Biochemistry* **30**, 4480–4486 (1991).