

Combination chemical genetics

Joseph Lehár^{1,2}, Brent R Stockwell³, Guri Giaever⁴⁻⁶ & Corey Nislow^{4,6,7}

Predicting the behavior of living organisms is an enormous challenge given their vast complexity. Efforts to model biological systems require large datasets generated by physical binding experiments and perturbation studies. Genetic perturbations have proven important and are greatly facilitated by the advent of comprehensive mutant libraries in model organisms. Small-molecule chemical perturbagens provide a complementary approach, especially for systems that lack mutant libraries, and can easily probe the function of essential genes. Though single chemical or genetic perturbations provide crucial information associating individual components (for example, genes, proteins or small molecules) with pathways or phenotypes, functional relationships between pathways and modules of components are most effectively obtained from combined perturbation experiments. Here we review the current state of and discuss some future directions for 'combination chemical genetics', the systematic application of multiple chemical or mixed chemical and genetic perturbations, both to gain insight into biological systems and to facilitate medical discoveries.

An important challenge facing the life sciences is to quantitatively describe the bewildering complexity of living organisms¹, both to appreciate the elegance of nature and to make medically relevant predictions. The scope of this complexity is vast. Even the function of a single mammalian cell typically involves coordinated activities among over 20,000 genes, 100,000 proteins², and thousands of small-molecule lipids, carbohydrates and metabolites, each of which may be expressed at differing levels over time. These components interact in physical complexes and functional modules that operate at many levels of organization¹. The systems biology approach to understanding such processes involves constructing large-scale models of cellular function, using networks of metabolic and signaling pathways extended and buttressed by incorporating interaction data obtained from physical binding experiments³.

Systems biology models are refined by exposing a system to experimental perturbations and assessing responses for consistency with model predictions³. A striking example of this approach was the construction of a detailed model of the pathways controlling sea urchin larval gut development that could predict specific abnormal phenotypes (for example, gut duplication) upon mutation⁴. Genetic perturbations are attractive because they make it possible to manipulate individual macromolecular components with little ambiguity about which target has been directly affected. Moreover, the advent of comprehensive

mutant libraries for several model organisms permits the large-scale investigation of phenotypic sensitivities across the entire genome, which can be extended through the use of complementary DNA and RNA interference libraries^{5,6}. However, genetic perturbations are labor intensive, cannot resolve protein functions that differ between contexts, and generate secondary effects by modulating the target protein's own abundance. Because chemical perturbagens avoid many of these concerns, genetic screens can be complemented by 'chemical genetics' studies that use organic molecules as perturbers⁷.

Although single perturbations are effective at determining which components in a system are essential for a phenotype, functional connections between components are best identified either by direct interaction data or through combination effects⁸. For example, paired genetic mutations can distinguish whether two nonessential genes have serial or parallel functionalities^{9,10}, and analyses across larger sets of perturbations can resolve the system into functional modules and pathways^{11,12}. With the arrival of genome-scale mutant libraries, comprehensive combination genetic experiments are now being undertaken in model organisms such as yeast⁸, nematodes¹³ and bacteria¹⁴. This approach has also been extended to chemical-gene interactions¹⁵⁻¹⁷ and purely chemical combinations^{18,19}, thus marking the advent of a new area of investigation: combination chemical genetics.

Combination chemical genetics (CCG) can be defined as the systematic testing of multiple perturbations involving chemical agents and can include purely chemical combinations or mixed chemical and genetic perturbations. Classical and chemical genetics are generally divided into 'forward' screens, in which uncharacterized perturbers are tested against a chosen phenotype to identify genes affecting that phenotype, and 'reverse' studies, in which a specific gene or protein is modulated and multiple phenotypes are monitored to determine the effects of that target^{7,20}. Studies involving combined perturbations can be similarly classified (Fig. 1) with the mechanistic focus shifted from individual targets to interactions between them. Here we present an overview of this emerging field and discuss uses of CCG for both scientific understanding and medical discoveries.

¹CombinatoRx Incorporated, 245 First Street, Cambridge, Massachusetts 02142, USA. ²Boston University Bioinformatics, 20 Cummington Street, Boston, Massachusetts 02215, USA. ³Columbia University, Department of Biological Sciences and Department of Chemistry, 614 Fairchild Center, MC2406, 1212 Amsterdam Avenue, New York, New York 10027, USA. ⁴Terrence Donnelly Center for Cellular and Biomolecular Research, University of Toronto, 160 College Street, Toronto, Ontario M5S 3E1, Canada. ⁵Department of Pharmaceutical Sciences, University of Toronto, 144 College Street, Toronto, Ontario M5S 3E1, Canada. ⁶Department of Molecular Genetics, University of Toronto, 1 King's College Circle, Room 4396, Toronto, Ontario M5S 1A8, Canada. ⁷Banting and Best Department of Medical Research, University of Toronto, 144 College Street, Toronto, Ontario M5S 3E1, Canada. Correspondence should be addressed to J.L. (jlehar@combinatorx.com).

Chemical genetics

Research in chemical genetics developed over time as a field derived from classical genetics, and most of the methods and terminology used reflect that history. Genetic knockouts have their counterparts in chemical 'knockdowns', and studies can be designed to be either forward or reverse depending on the direction of learning that underlies their motivation^{7,20}. Forward studies involve testing many chemical probes against one or a few phenotypes in order to identify active agents, and reverse studies perform multiple phenotype measurements on a few related chemicals to characterize their function. In both cases, chemical agents can be analyzed across a panel of phenotypic assays to identify either broad activity or selectivity between the phenotypes. The essential tools for chemical genetics include large, diverse libraries of chemicals with both known and unknown biological targets²¹ (Table 1). These can comprise approved drugs²² and mechanistically characterized chemical probes²³, which are often informed by chemical-protein associations from the literature²⁴.

Chemical perturbations provide information that is distinct from and complementary to the information provided by genetic mutations, given the differences between how they modulate protein functions⁷. The advantages of chemical perturbations are that they (i) can target a single domain of a multidomain protein, (ii) allow precise temporal control that is critical for rapidly acting processes, (iii) can target orthologous or paralogous proteins, enabling comparisons between species or redundant functions, and (iv) do not directly alter the concentrations of a targeted protein, thus avoiding indirect effects on multiprotein complexes. Small molecules also lend themselves more readily to combination interventions, making them especially useful for integrating systems and chemical biology^{9,19}.

The main challenges with using small molecules are that they are generally pleiotropic—they have multiple dose-dependent molecular targets that are often not fully characterized, which leads to unexpected activities. Some small molecules also have additional liabilities that can hamper their use, such as *in vitro* aggregation, poor solubility, difficulty traversing biological membranes and reactive or toxic functionalities. Moreover, compounds are subject to metabolic modifications *in vivo* that can substantially alter their activities and toxicities. Such liabilities can be minimized in many cases, but suitable precautions need to be taken; for example, multiple structurally distinct probes sharing a known target can be used to distinguish off-target from on-target effects. Finally, despite the impressive size of chemical databases (Table 1), the known targets of bioactive libraries still cover only a small fraction of the proteome²². This deficiency is due to the difficulty of finding biological targets, constraints such as cellular compartmentalization or varying protein levels that make some targets inaccessible, and the limited structural diversity of libraries generated by combinatorial chemistry²⁵.

Forward chemical genetic approaches have recently yielded medically relevant and biologically informative insights. For example, comparing activity profiles across ~70,000 compounds revealed that modulating the activity of mitochondrial voltage-dependent anion channels causes the selective death of RAS-transformed tumor cells²⁶, and subsequent mechanistic studies²⁷ showed that three such compounds achieve this selectivity through a new, non-apoptotic cell death process. Similarly, gene expression profiles in cancer cell lines identified genetic markers for acute myelogenous leukemia that were then used for a chemical genetic screen across a diverse set of 1,739 compounds, to find 8 drug candidates that induced favorable cell differentiation signatures²⁸. As a final example, a cell-based screen of small-molecule libraries identified chemicals that induce stem cell self renewal through modulating specific combinations of targets²⁹.

Reverse chemical genetics studies have been similarly revealing. For example, a number of annotated compounds were screened for their ability to prevent the enigmatic cell death process that occurs in olfactory sen-

sory neurons, among the few types of neurons that turn over *in vivo*³⁰. The study found that jun kinase (JNK) inhibitors could prevent this cell death by means of a new functional role for the target, and it validated this new signaling function for JNK both *in vitro* and *in vivo*. As another example, after discovering inhibitors of the yeast transcription factor component Hap3p in forward screens using immobilized small-molecule arrays, the new inhibitor haptamide B was mechanistically characterized using whole-genome transcriptional profiling of wild-type and knockout yeast cells³¹. Finally, a compendium of expression profiles was used to functionally characterize small-molecule treatments³², which demonstrates that transcriptional profiling is one of the most effective tools for reverse chemical genetics.

Experiments that test many compounds against a comparable number of phenotypes can be used to integrate both forward and reverse approaches in a single study. For example, 100 diverse drugs and research

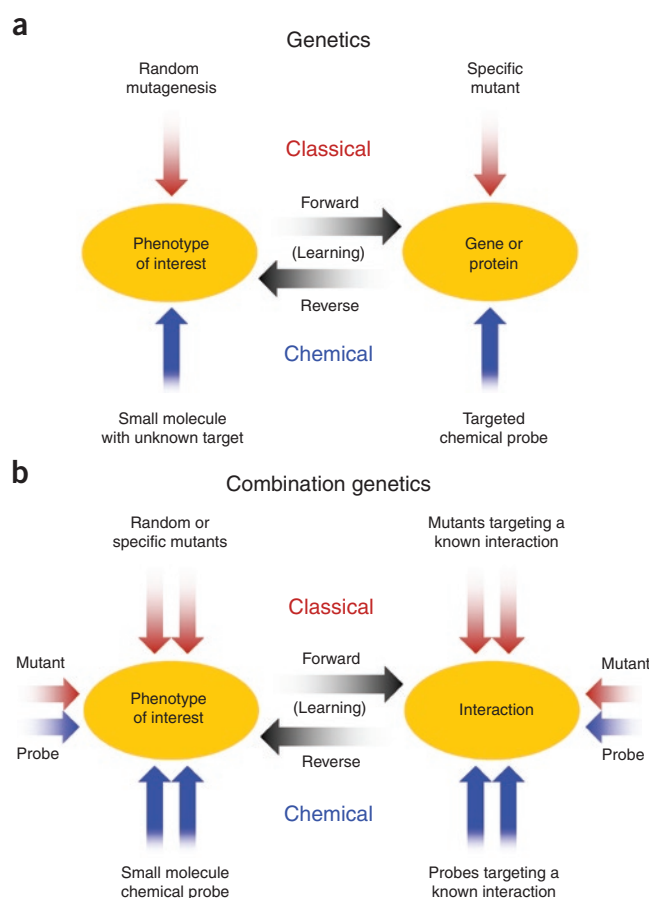


Figure 1 Combined perturber studies in the context of forward and reverse genetics. (a) In classical and chemical genetics, “forward” screens use many uncharacterized perturbors and a known phenotype to discover genes or proteins that affect that phenotype, and “reverse” studies test many phenotypes using a perturbagen with a known target to determine which phenotypes are affected by that target. In both cases, the questions under investigation center on the function of individual genes or proteins. (b) For combination chemical genetics, the focus of investigations shifts from individual targets to interactions between them or conditional target dependencies, and the perturbations are applied as combinations. Here forward screens use combinations of many perturbagens to discover interactions, and reverse studies involve modulating a known interaction with a set of probes targeting its components to determine which of many tested phenotypes are affected by that interaction. Figure adapted from ref. 20.

Table 1 Perturbation sets for combination chemical genetics

Chemical sets	Number of probes	Number of targets	Notes
Existing drugs	1,357	324–1,048	Drugs approved in the United States ²²
Drugs and probes	4,765	4,447	DrugBank ²³ (www.drugbank.ca/ , on 2008.08.18)
Bioactives	37,349	6,128	PubChem actives ¹⁰⁰ (http://pubchem.ncbi.nlm.nih.gov/)
Literature	128,120	1,320	WOMBAT ^{24,100} (version 2006.1)
Registered chemicals	$>3.7 \times 10^7$	–	CAS ¹⁰¹ (http://www.cas.org/ , on 2008.08.18)
Genetic libraries	Number of probes	% of genome	Notes
Bacteria, <i>E. coli</i>	~3,900	93%	KO ⁷⁷ , ORF ⁷⁸
Bacteria, <i>S. aureus</i>	~2,600	95%	ORF ⁷⁹
Bacteria, <i>Brucella</i>	~3,000	90%	ORF ⁸⁰
Bacteria, <i>Campylobacter</i>	~1,600	98%	ORF ⁸¹
Bacteria, <i>F. tularensis</i>	~1,500	96%	ORF ⁸²
Fungi, <i>S. cerevisiae</i>	~6,100	98%	KO ¹⁵ , OX ⁸⁹ , ORF ¹⁰²
Fungi, <i>S. pombe</i>	~5,000	95%	KO (http://www.bioneer.com/), ORF ⁸³
Fungi, <i>C. albicans</i>	~2,800	45%	KO ⁸⁴
Fungi, <i>C. glabrata</i>	~500	10%	KO ⁸⁵
Worm, <i>C. elegans</i>	~11,000	50%	RNAi ⁸⁷
Fly, <i>Drosophila</i>	~13,000	95%	RNAi ⁸⁸ , ORF (http://www.fruitfly.org/)
Vertebrate, <i>Mus musculus</i>	~11,000	50%	RNAi ⁶ , shRNA ⁵
Vertebrate, <i>Homo sapiens</i>	~22,000	90%	RNAi ⁶ , ORF ¹⁰³

KO, knockout; OX, overexpression; ORF, open reading frame; RNAi/shRNA, RNA interference.

probes were tested in cancer cells across ~100 phenotypes (for 11 cancer-relevant proteins, each with ~10 microscopic measurements of stains or fluorescent antibodies tested at 12 concentrations of each drug)³³, and the resulting dataset was used both for assigning targets to previously uncharacterized drugs and for measuring the functional responses to drugs with remarkable detail. Also in this category is the Connectivity Map project³⁴, which is systematically collecting whole-genome expression profiles across multiple tumor cell lines for a large number of chemical agents, with both a forward goal of identifying chemicals with desirable selectivity across a multigene expression phenotype and a reverse objective of mechanistically characterizing drug responses.

Combined perturbors

Combined perturbation studies are inherently more complex than those based on single agents. In particular, the interaction needs to be compared to the individual single-agent effects in order to determine whether there is “synergy,” where the agents cooperate toward a phenotype, or “antagonism,” where they impede each other’s activity. Systematic combination experiments also require quantitative models that represent the expected combination effects, against which synergy and antagonism can be measured.

Historically, interactions between genes have been described as epistatic relationships deriving from statistical concepts³⁵. In these models, the fitness of a double mutant in the absence of a genetic interaction is expected to be the product of the individual fitness measurements of the corresponding single mutants. Another definition of epistasis derives from the work of Bateson and is typically used to describe situations in which the activity of one gene masks effects at another locus, allowing inferences about the order of gene action. Classical examples of Bateson-type epistasis analysis include studies of signaling pathways that control the yeast cell cycle, nematode pheromone responses and sex determination in *Drosophila melanogaster*³⁵. More recently, defining the nature of genetic interactions has been expanded^{36,37} and modeled in the context of metabolic¹⁰ or

other cellular networks³⁸.

When working with chemicals, the effects of varying the concentrations need to be considered³⁹ (Fig. 2). Chemical agents generally transition from inactive to a fixed effect level as the dose increases, and the rate of transition depends on the mechanistic interaction between the perturber and its target (for example, whether the chemical is a competitive or allosteric inhibitor). Introducing a second agent can cause synergy or antagonism either by boosting the high-dose response to a different effect level or by shifting the transition to a higher or lower concentration. Many chemical perturbors also affect multiple targets, in some cases at distinct doses, leading to several transition levels in their response curves and further complicating any combination analysis. Chemical responses can be measured for drug-gene interactions using serially diluted dose series experiments with and without the genetic perturbation. For chemical combinations, however, the optimal dosing ratio between the agents is unknown, so it is useful either to test the combination as a fixed dose-ratio series, where component drugs are mixed at a high concentration and the mixture is diluted serially, or to use a factorial “dose matrix” testing all pairs of serially diluted single-agent doses. The results from a dose matrix experiment can be visualized as a three-dimensional response surface, or from the top using color to represent the different effect levels. By focusing on a chosen effect contour with an “isobologram,” it is possible to visualize how much dose shifting has occurred at that level owing to synergy or antagonism between the agents.

Determining synergy for chemical combinations is also more complicated³⁹. The genetic interaction models have their counterparts (Fig. 2) in Bliss independence (multiplicativity) and Gaddum’s non-interaction (masking). For medical applications, however, the most relevant reference is Loewe additivity, which is the dose-additive expectation for a drug combined with itself, usually compared to observed synergies via a combination index⁴⁰. Although all three reference models are indistinguishable when one of the perturbations has no effect alone, there has been considerable disagreement concerning which null-interaction model is most generally applicable to biological contexts^{19,37,39}. For now, it seems that genetic interactions in proliferation assays find mostly multiplicative

effects³⁷, whereas chemical combinations are dominated by masking responses¹⁹. Ultimately, comparisons between genetic and chemical effects should be made using a single null-effect reference that is mechanistically appropriate to a common measured endpoint.

Going beyond simply measuring synergy or antagonism, dose matrix responses to chemical combinations can also provide detailed mechanistic information. Although modeling combination effects for theoretical networks of reactions has historically proven difficult⁴¹, largely because the combination index at a chosen effect level can be very sensitive to the assumed kinetic parameters in each reaction, more recent efforts focused on the dynamic responses of multiply inhibited simulated networks across different effect levels have yielded more stable predictions^{19,42}. When performed in dose matrices, such simulations produce a variety of synergistic and antagonistic responses whose surface shapes contain information on the topological connectivity between inhibited targets¹⁹ (Fig. 3). Experimental dose matrices can be morphologically compared to responses from many simulations like these, each of which represents a mechanistic hypothesis about how the targets are connected in the network under study, to determine which hypotheses are most consistent with the data.

The experimental designs for CCG studies bear a direct resemblance to those used for classical and chemical genetics (Fig. 4). Tests in each phenotype become either paired-dose series or full-dose matrices, depending on whether the perturbations can be continuously dosed, and in each case the combination data are used to derive a score quantifying synergy or antagonism between the perturbations. By analogy with chemical genetics, forward CCG studies involve testing multiple combinations against one or a few phenotypes in order to identify synergistic interactions or conditional sensitivities, and reverse studies perform multiple phenotype measurements on combinations aimed at a few related interactions to characterize their function. Both approaches can be integrated by collecting profiles across many combinations for many phenotypes, thus enabling large-scale profile comparisons in either direction.

Chemogenomics

The most advanced area of combined chemical genetics involves the joint application of chemical and genetic perturbations. The budding yeast *Saccharomyces cerevisiae* has dominated this field, owing to the comprehensive genetic tools that have been developed. These resources include a systematic gene deletion set, where precise start-to-stop mutant strains were constructed for ~6,000 genes (homozygous for inessential and heterozygous for essential genes)¹⁵. The strains were constructed with molecular tags to permit individual mutant sensitivities to be extracted

from genome-wide competitive growth in a single culture⁴³. Mutant libraries are being developed for a number of other model organisms, and the construction of large RNA interference (RNAi) probe sets and open reading frame (ORF) libraries (Table 1) has opened up still more organisms for systematic genetic testing. Using these tools, “chemogenomic” chemical sensitivity screens can be carried out that can test large sets of cells with mutated, silenced or overexpressed genes against panels of chemical perturbations.

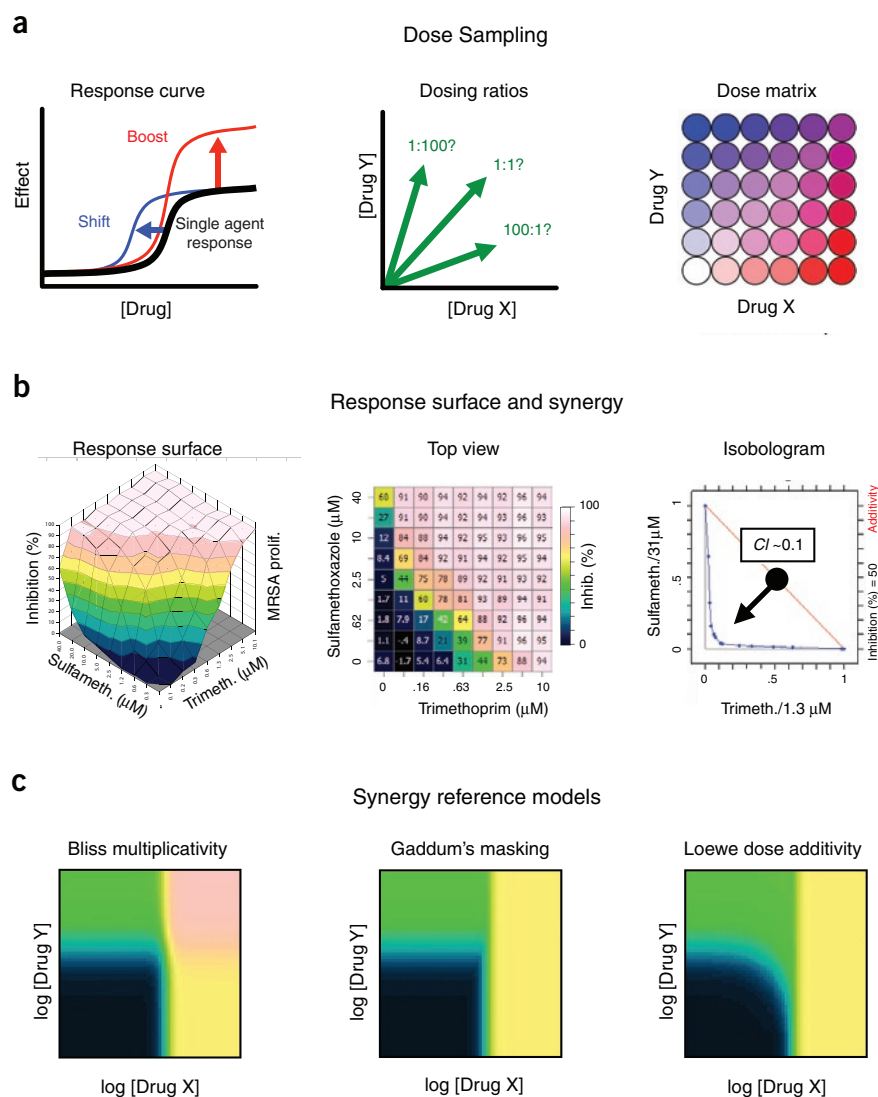


Figure 2 Measuring synergy for chemical combinations. (a) Continuous perturbations with sigmoidal response curves can cooperate either to boost the high-dose effect to new levels or to shift the effective concentration to lower doses, and the optimal dosing ratio is usually not known. A factorial dose matrix design can capture all of these possibilities. (b) The resulting interaction can be analyzed using the full three-dimensional response surface or using an isobologram to measure linear dose shifting at a chosen effect level via a combination index CI . For this example, we show a strongly synergistic antibiotic combination that targets folate metabolism enzymes¹⁹. (c) Synergy reference models will differ depending on the null-effect assumption. Bliss independence (multiplicative epistasis) or Gaddum's non-interaction (Bateson masking) are generally used to analyze genetic epistasis, and Loewe dose additivity is most widely used for drug combinations. The multiplicative model produces stronger effects than either of the single agents at high combined doses, whereas masking simply follows the strongest single agent at corresponding doses. In dose-additive combinations, the agents cooperate in the same way as increasing the dose of a single drug. All three models can be adapted for analyzing pairs of agonists and generalized for three or more agents.

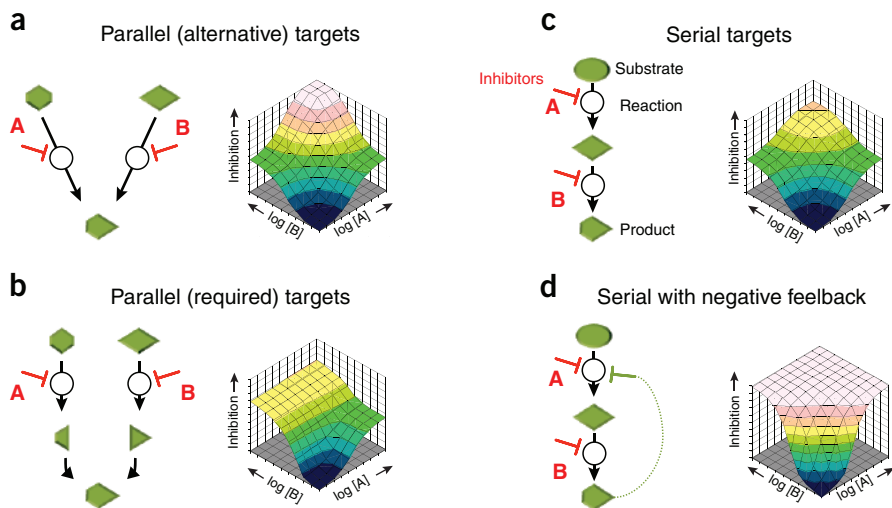


Figure 3 The response shape in dose matrix experiments depends on target connectivity. In simulations of multiply inhibited metabolic networks¹⁹, the response surface morphology depended strongly on how the inhibited targets were connected in the network. Here we show four representative target connectivities, where substrates (green symbols) are metabolized by reactions (black arrows), and the reaction enzymes (white circles) are modulated by inhibitors (red markers). The resulting response surfaces from dynamic simulations are shown to the right of each such pathway. (a,b) Inhibitor pairs with parallel targets produced either saturated (a) or masking (b) effects in combination, depending on whether the targets affected independent alternatives or codependent ingredients of the final reaction. (c,d) Inhibiting serial targets along a pathway yielded multiplicative effects (c) for partially effective single agents, but serial targets in pathways regulated by negative feedback (d) produced strong dose shifting like that seen in our antibacterial example (Fig. 2b). Each of these cases represents a mechanistic hypothesis relating response shape to target connectivity, generated by the pathway simulations.

Early large-scale efforts in CCG focused on discovering interactions between chemical and genetic perturbations in bacteria, beginning with statistical epistasis between random mutagenesis and chemical stresses⁴⁴. The introduction of large-scale yeast mutant libraries, however, caused a rapid transition to chemogenomic screening, which had both the forward and reverse aspects of integrated chemical genetics. For example, a screen of 74 drugs and chemical probes against a panel of ~3,000 heterozygous yeast mutants identified sensitive strains as likely targets for most of the chemicals¹⁷. Another study that tested ten drugs at multiple doses against a comprehensive panel of yeast mutants identified multiple genes that conferred sensitivity to each of the drugs¹⁶. Mutant sensitivities involving haploinsufficient essential genes were used to identify direct targets of the drug, but those from homozygous deletions discovered gene products that act indirectly on the drug's targets¹⁶. When such drug sensitivity profiles are compared with double-mutant fitness profiles involving a drug target's pathway⁴⁵, the degree of similarity can provide further information on the drug's mechanism of action. As an example of reverse CCG among these studies, chemogenomic screening with homozygous strains produced insights on the machinery of RNA processing¹⁶. Chemogenomic screens in yeast cells have also been used to discover a new antineoplastic mechanism that induces mitotic arrest⁴⁶ and to identify potential toxicities related to the off-target effects of psychoactive drugs⁴⁷.

Chemogenomic screening in other organisms has also advanced, often with a more direct focus on drug discovery applications. For example, screening natural compounds using a *Candida albicans* mutant set identified fungal-specific inhibitors of mRNA polyadenylation⁴⁸, thereby opening a new mechanistic class for potential antifungal therapies. In human cells, combining RNAi perturbation with small molecules⁴⁹ offers insights into the activity of chemical combinations on both normal and abnormal disease-relevant cell types. For example,

small interfering RNA screens with commonly used antineoplastics in tumor cell lines have identified potential targets for cotherapeutic cancer treatment⁵⁰. Whole-genome expression signatures have also been used as a phenotype to identify a beneficial interaction between the mTOR inhibitor rapamycin and the glucocorticoid receptor as a potential therapy for lymphoid malignancies⁵¹.

As the number of genetic perturber sets expands (Table 1), chemogenomic approaches will provide ever larger sets of interaction data to aid with systems biology modeling efforts⁵². The analysis of chemogenomic screens presents a formidable challenge because they involve profiles with thousands of combinations, where each single agent might have multiple states of genetic perturbations (knockout, knockdown, over-expression) and many doses of chemical components. Such datasets will require both appropriate synergy analyses for each individual response matrix¹⁹ and global displays and statistics for determining mechanistic patterns between the interactions⁵³.

Combinations of chemicals

The usefulness of synergistic drug combinations has long been appreciated^{19,39,54,55}, especially in the therapeutic areas of cancer and infectious diseases. Multitarget synergies between drugs

can be used to overcome resistance to one of the components, to make use of targets that are individually insufficient for therapy and to reduce side effects through dose sparing⁵⁶. However, although there have been numerous individual studies of synergistic combinations³⁹, it is only recently that systematic large-scale testing of chemical combinations has been pursued^{18,19,57}. Owing to the required investment in concentration sampling, these efforts have focused on testing integrative phenotypes, such as growth, that are influenced by many cellular components, rather than monitoring many genes simultaneously.

The first systematic CCG efforts involving pure chemical combinations were aimed directly at drug discovery. Using a dedicated screening platform, an effort focused on cancer, inflammation and fungal assays identified a number of synergies as potential combination therapeutics⁵⁷. All three assays yielded unexpected interactions involving drugs not normally administered toward the target therapy. Both the anticancer and the anti-inflammation synergy translated to animals⁵⁷, and the latter has since been confirmed in a phase 2a clinical trial⁵⁸. Drug discovery approaches of this kind must consider the possibility of synergistic toxicity and *in vivo* pharmacological interactions⁵⁹ that can complicate delivery of the combination to targeted tissues or cause metabolic hepatotoxicity. Synergies in cell-based assays can also fail to translate when tested in animals or humans. To avoid such concerns, it is essential to carry out counterscreens and preclinical investigations, which can be guided by published toxicity or drug interaction data for the single agents, before considering any potential combination therapies for the clinic.

The use of chemical combination screens to extract mechanistic information is at an earlier stage. The first such screen tested all pairs of 21 antibiotic drugs in *Escherichia coli* bacterial proliferation at single combination dose points¹⁸. This study found a wide variety of combination effects between the drugs and was able to identify modules of drug

pairs whose group interactions were purely synergistic or antagonistic, where each module comprised drugs targeting related cellular functions. Another chemical genetic screen in *Candida glabrata* yeast¹⁹, testing all pairs of ten antifungal drugs in six-by-six dose matrices, confirmed the relationship between response matrix morphology and target connectivity expected from dynamic pathway simulations¹⁹ (Fig. 3). Also in that study, an analysis of synergy profiles derived from an anticancer screen in HCT116 cells that tested all pairs of 90 drugs and research probes in six-by-six dose matrices found that probes with related targets had more similar synergy profiles than those with disparate functions¹⁹. These examples illustrate how CCG using chemical combinations, especially with varying doses, can reveal constraints on the topology of the underlying cellular network and assist with identifying unknown activities of chemical probes in the library. It is important to note that the relevance of combination effects from phenotypic experiments is limited to the cell systems under study and that mechanistic conclusions from *in vitro* synergies can fail to translate to other contexts (for example, due to cell type differences or *in vivo* pharmacological drug interactions). In any case, such studies are dependent on libraries of well-characterized chemical probes with known biological targets, and it is especially valuable to have several chemically distinct probes targeting the same protein when possible in order to separate on-target (consistent responses) from off-target (inconsistent response) effects.

Almost all of the past CCG screens involving combinations of chemicals have been essentially forward studies, aimed at uncovering unexpected synergies or interactions between drug targets. Reverse CCG studies are a more recent development but are increasingly undertaken, especially to elucidate the mechanisms of anticancer synergies. For example, whole-genome transcriptional profiling in prostate cancer cell lines of synergies between taxane microtubule binders and either estramustine⁶⁰ or capecitabine⁶¹ identified genes associated with the synergistic response that were not present in the single agents' profiles. Another study testing the combination of the apoptosis inducer taurilidine with tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) used esophageal carcinoma cells and flow cytometry to characterize the synergistic responses in many apoptosis-related signaling proteins⁶². In yeast, chemogenomic screening performed over multiple doses of several drug combinations was used to map strain-specific interaction measurements onto the yeast genetic network and identify protein complexes that buffer the cell from the drug combination⁵³. Given recent trends, we expect such studies to increase in frequency and complexity.

Discussion and future directions

Combination chemical genetics is an emerging field of research. Whether genetic and chemical perturbations are applied together or combinations are purely between chemicals, CCG allows the testing of interactions between cellular components to be studied in new contexts and with more detail than can be achieved with single agents or only genetic perturbations.

To make the best use of available resources, CCG researchers will need to agree on standards for the design, data collection and analysis of experiments, building on resources already established for chemical genetics. For example, the US National Institutes of Health (NIH) Molecular Library Initiative⁶³ aims to create a publically available collection of biologically active compounds⁶⁴ and experimental standards⁶⁵. For CCG, it will be helpful to coordinate common probe libraries that can facilitate comparisons between experiments, and to agree on combination effect reference models for analyses across combination studies. Interaction data can be collected in existing public repositories (for example, BioGRID⁶⁶), which will need to be equipped with cheminformatics tools to allow integration with the knowledge in chemical databases²¹.

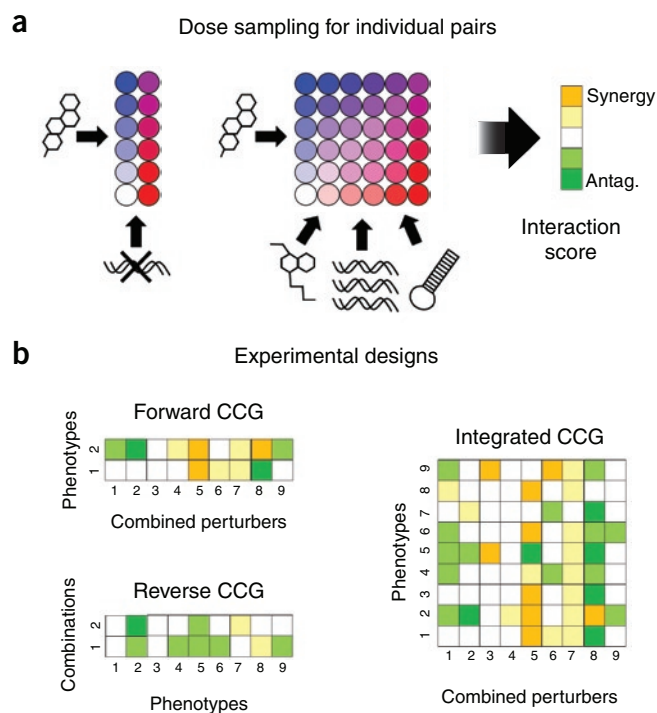


Figure 4 Designing combination chemical genetics experiments. **(a)** The dose sampling possible for individual pairs depends on whether the perturbagens are discrete (for example, knockouts) or continuous (for example, chemicals, overexpression or RNAi). When the perturbers have known cellular targets, interactions can be described in terms of those targets. **(b)** Following the practices of chemical genetics, CCG involves many such experiments, either testing multiple combinations against a few phenotypes to discover synergistic interactions (forward) or testing a few combinations against many phenotypes to characterize the function of an interaction (reverse). These approaches can be integrated by collecting profiles across many combinations for a comparable number of phenotypes, allowing profiles in either direction to be compared for similarity or selective effects.

Chemical genetic studies will benefit from continued improvement and extension of available chemical libraries. Efforts to catalog the druggable genome for existing chemicals²² have led to collections of biologically active compounds with known targets^{23,67}. Ligands for protein targets can be identified using traditional affinity purification⁷, radiolabeling⁶⁸ and live-cell target-labeling approaches⁶⁹. Yeast three-hybrid techniques⁷⁰ and the use of immobilized chemical or protein arrays⁷¹ can reveal direct drug-protein binding events, and haploinsufficiency fitness tests¹⁵ can associate small molecules with their target genes, which will be especially helpful toward characterizing the mechanisms of bioactive natural products⁷². Chemical genetics probe sets are also being extended by comparing response profiles across cell-based phenotypes^{33,67}, by gene expression profiling^{28,34} and by analysis of drug activities and side effects⁷³. Considerable progress has also been made with establishing libraries for specific target classes (for example, epidermal growth factor receptor kinases⁷⁴). For modulating the many targets that are not covered by current probe sets, it remains critical to develop libraries with greater chemical diversity. Current efforts to this end involve assembling bioactive molecular fragments⁷⁵, or using diversity-oriented synthesis⁷⁶ to produce complex, natural product-like libraries.

Chemogenomics efforts are becoming increasingly complex and diverse with the introduction of expanded probe sets and higher content experimental platforms. The resources for genome-scale genetic perturbations

in *S. cerevisiae* and other yeasts have grown dramatically, and the set of model organisms has expanded to include bacteria^{77–82}, other yeasts^{83–85}, invertebrates^{13,86–88} and select vertebrates^{5,6} (Table 1). Genome-wide overexpression libraries are being constructed in yeast that can be applied in conjunction with knockout experiments to provide a complete characterization of genetic perturbations⁸⁹. Finally, the rapid adoption of RNAi technology promises a smooth transition from sequence to function in more complex cells and metazoans⁹⁰, and flow cytometry readouts permit high-throughput analysis of changes in cell-cycle state or the expression of cytoplasmic and cell-surface markers.

Drug discovery will remain a major driving force for combination chemical genetics, and we expect to see expanded efforts involving forward, reverse and integrated CCG approaches toward this goal. One advantage of forward CCG is that it is a purely empirical approach, allowing new biological interactions to be revealed. Another potential advantage is suggested by a recent comparison of synthetic lethality in yeast (with deletion alleles) and nematodes (using double RNAi), which concluded that synthetic lethal interactions are not conserved⁹¹. This suggests that the kinds of interactions probed by CCG are likely to be organism and context specific, offering the possibility that combination therapies targeting such interactions may achieve higher levels of selectivity than single agents toward targeting infectious diseases or other context-dependent conditions such as cancer.

Considering theoretical simulations, an obvious extension of the current modeling is to simulate chemical combination effects at genome scales. This can be achieved by adapting the metabolic simulations¹⁰ previously used for genetic interaction predictions to allow partial inhibition rather than total knockouts for target genes. The response surfaces from such simulations could be compared to data from matched combination experiments, and inconsistencies could guide improvements in our understanding of the underlying biological network. Another promising approach is to use combination effects to infer topological models of the target network, using linear or nonlinear regression methods⁹² adapted for combination data constraints⁹³. This approach provides a data-driven complement to the model-driven predictions from a priori reaction networks, and combining both methods should provide an effective strategy for refining predictive biological models.

An area with considerable promise involves high-order combinations of three or more perturbagens. Current CCG studies in yeast extend to third order (combinations of three agents), using designs that (i) test pairwise combinations against single mutants⁹⁴ or chemogenomic profiles⁵³, (ii) screen double mutants for sensitivity to drug treatments⁹⁵ or varying conditions⁹⁶ and (iii) investigate purely chemical combinations⁹⁷. All three designs generally find synergies that pairwise interactions could not fully account for. Theoretically, high-order combinations should yield ever more selective control of complex systems⁴², and it should even be possible to use high-order testing to quantify the functional complexity of a biological system⁹⁸. Extending chemical genetic studies to yet higher order combinations should provide constraints on the limits of medically useful synergy⁹⁸ as well as mechanistic insights into biological networks⁹⁹. Another dimension to explore is the effect on synergy due to nonsimultaneous drug application, both on heterogeneous and synchronized cell populations, which could reveal conditional dependencies on cellular state changes. Designing and analyzing high-order and phased combination experiments will be challenging, and it will be an area of considerable activity in coming years.

Combination chemical genetics brings together the traditions of genetic perturbation and synergistic drug discovery to enable the detailed study of network topology. This has been made possible by the assembly of large, diverse chemical libraries and comprehensive sets of genetic mutants or RNAi suppressors. This marriage of large-scale genomic

approaches, synergy analysis and chemical genetic tools offers the promise of new insights into biology and a new avenue for drug discovery.

ACKNOWLEDGMENTS

B.R.S. is supported by a Beckman Young Investigator Award from the Arnold and Mabel Beckman Foundation and by the NIH (CA097061 and GM085081). C.N. and G.G. are supported by the Canadian Institutes of Health Research (CIHR MOP-81340 and CIHR MOP-84305, respectively).

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at <http://www.nature.com/naturechemicalbiology/>.

Published online at <http://www.nature.com/naturechemicalbiology/>

Reprints and permissions information is available online at <http://ngp.nature.com/reprintsandpermissions/>

- Stelling, J., Sauer, U., Szallasi, Z., Doyle, F.J. III & Doyle, J. Robustness of cellular functions. *Cell* **118**, 675–685 (2004).
- Koonin, E.V., Wolf, Y.I. & Karev, G.P. The structure of the protein universe and genome evolution. *Nature* **420**, 218–223 (2002).
- Hood, L., Heath, J.R., Phelps, M.E. & Lin, B. Systems biology and new technologies enable predictive and preventative medicine. *Science* **306**, 640–643 (2004).
- Davidson, E.H. *et al.* A genomic regulatory network for development. *Science* **295**, 1669–1678 (2002).
- Silva, J.M. *et al.* Second-generation shRNA libraries covering the mouse and human genomes. *Nat. Genet.* **37**, 1281–1288 (2005).
- Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**, 1283–1298 (2006).
- Stockwell, B.R. Exploring biology with small organic molecules. *Nature* **432**, 846–854 (2004).
- Boone, C., Bussey, H. & Andrews, B.J. Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* **8**, 437–449 (2007).
- Sharom, J.R., Bellows, D.S. & Tyers, M. From large networks to small molecules. *Curr. Opin. Chem. Biol.* **8**, 81–90 (2004).
- Segrè, D., Deluna, A., Church, G.M. & Kishony, R. Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83 (2005).
- Wong, S.L. *et al.* Combining biological networks to predict genetic interactions. *Proc. Natl. Acad. Sci. USA* **101**, 15682–15687 (2004).
- Zhang, L.V. *et al.* Motifs, themes and thematic maps of an integrated *Saccharomyces cerevisiae* interaction network. *J. Biol.* **4**, 6 (2005).
- Kamath, R.S. *et al.* Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237 (2003).
- Butland, G. *et al.* eSGA: *E. coli* synthetic genetic array analysis. *Nat. Methods* **5**, 789–795 (2008).
- Giaever, G. *et al.* Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391 (2002).
- Giaever, G. *et al.* Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc. Natl. Acad. Sci. USA* **101**, 793–798 (2004).
- Lum, P.Y. *et al.* Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell* **116**, 121–137 (2004).
- Yeh, P., Tschumi, A.I. & Kishony, R. Functional classification of drugs by properties of their pairwise interactions. *Nat. Genet.* **38**, 489–494 (2006).
- Lehár, J. *et al.* Chemical combination effects predict connectivity in biological systems. *Mol. Syst. Biol.* **3**, 80 (2007).
- Kawasumi, M. & Nghiem, P. Chemical genetics: elucidating biological systems with small-molecule compounds. *J. Invest. Dermatol.* **127**, 1577–1584 (2007).
- Oprea, T.I., Tropsha, A., Faulon, J.L. & Rintoul, M.D. Systems chemical biology. *Nat. Chem. Biol.* **3**, 447–450 (2007).
- Overington, J.P., Al-Lazikani, B. & Hopkins, A.L. How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 993–996 (2006).
- Wishart, D.S. *et al.* DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **36**, D901–D906 (2008).
- Olah, M. *et al.* in *Chemical Biology: from Small Molecules to Systems Biology and Drug Design* Vol. 2 (eds. Schreiber, S.L., Kapoor, T. & Wess, G.) 760–779 (Wiley-VCH GmbH, Weinheim, Germany, 2007).
- Feher, M. & Schmidt, J.M. Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* **43**, 218–227 (2003).
- Yagoda, N. *et al.* RAS-RAF-MEK-dependent oxidative cell death involving voltage-dependent anion channels. *Nature* **447**, 864–868 (2007).
- Yang, W.S. & Stockwell, B.R. Synthetic lethal screening identifies compounds activating iron-dependent, nonapoptotic cell death in oncogenic-RAS-harboring cancer cells. *Chem. Biol.* **15**, 234–245 (2008).
- Stegmaier, K. *et al.* Gene expression-based high-throughput screening (GE-HTS) and application to leukemia differentiation. *Nat. Genet.* **36**, 257–263 (2004).
- Chen, S. *et al.* Self-renewal of embryonic stem cells by a small molecule. *Proc. Natl. Acad. Sci. USA* **103**, 17266–17271 (2006).
- Gangadhar, N.M., Firestein, S.J. & Stockwell, B.R. A novel role for jun N-terminal kinase signaling in olfactory sensory neuronal death. *Mol. Cell. Neurosci.* **38**, 518–525

- (2008).
31. Koehler, A.N., Shamji, A.F. & Schreiber, S.L. Discovery of an inhibitor of a transcription factor using small molecule microarrays and diversity-oriented synthesis. *J. Am. Chem. Soc.* **125**, 8420–8421 (2003).
 32. Hughes, T.R. *et al.* Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126 (2000).
 33. Perlman, Z.E. *et al.* Multidimensional drug profiling by automated microscopy. *Science* **306**, 1194–1198 (2004).
 34. Lamb, J. *et al.* The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
 35. Avery, L. & Wasserman, S. Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends Genet.* **8**, 312–316 (1992).
 36. Anastassiou, D. Computational analysis of the synergy among multiple interacting genes. *Mol. Syst. Biol.* **3**, 83 (2007).
 37. Mani, R., St Onge, R.P., Hartman, J.L. IV, Giaever, G. & Roth, F.P. Defining genetic interaction. *Proc. Natl. Acad. Sci. USA* **105**, 3461–3466 (2008).
 38. Carter, G.W. *et al.* Prediction of phenotype and gene expression for combinations of mutations. *Mol. Syst. Biol.* **3**, 96 (2007).
 39. Greco, W.R., Bravo, G. & Parsons, J.C. The search for synergy: a critical review from a response surface perspective. *Pharmacol. Rev.* **47**, 331–385 (1995).
 40. Chou, T.C. & Talalay, P. Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv. Enzyme Regul.* **22**, 27–55 (1984).
 41. Jackson, R.C. Amphibolic drug combinations: the design of selective antimetabolite protocols based upon the kinetic properties of multienzyme systems. *Cancer Res.* **53**, 3998–4003 (1993).
 42. Araujo, R.P., Petricoin, E.F. & Liotta, L.A. A mathematical model of combination therapy using the EGFR signaling network. *Biosystems* **80**, 57–69 (2005).
 43. Costanzo, M., Giaever, G., Nislow, C. & Andrews, B. Experimental approaches to identify genetic networks. *Curr. Opin. Biotechnol.* **17**, 472–480 (2006).
 44. Elena, S.F. & Lenski, R.E. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**, 395–398 (1997).
 45. Parsons, A.B. *et al.* Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* **126**, 611–625 (2006).
 46. Dorer, R.K. *et al.* A small-molecule inhibitor of Mps1 blocks the spindle-checkpoint response to a lack of tension on mitotic chromosomes. *Curr. Biol.* **15**, 1070–1076 (2005).
 47. Ericson, E. *et al.* Off-target effects of psychoactive drugs revealed by genome-wide assays in yeast. *PLoS Genet.* **4**, e1000151 (2008).
 48. Jiang, B. *et al.* PAP inhibitor with in vivo efficacy identified by *Candida albicans* genetic profiling of natural products. *Chem. Biol.* **15**, 363–374 (2008).
 49. Eggert, U.S. *et al.* Parallel chemical genetic and genome-wide RNAi screens identify cytokinesis inhibitors and targets. *PLoS Biol.* **2**, e379 (2004).
 50. MacKeigan, J.P., Murphy, L.O. & Blenis, J. Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nat. Cell Biol.* **7**, 591–600 (2005).
 51. Wei, G. *et al.* Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. *Cancer Cell* **10**, 331–342 (2006).
 52. Hillenmeyer, M.E. *et al.* The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* **320**, 362–365 (2008).
 53. Hoon, S. *et al.* An integrated platform of genomic assays reveals small-molecule bioactivities. *Nat. Chem. Biol.* **4**, 498–506 (2008).
 54. Keith, C.T., Borisy, A.A. & Stockwell, B.R. Multicomponent therapeutics for networked systems. *Nat. Rev. Drug Discov.* **4**, 71–78 (2005).
 55. Kaelin, W.G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat. Rev. Cancer* **5**, 689–698 (2005).
 56. Zimmermann, G.R., Lehár, J. & Keith, C.T. Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov. Today* **12**, 34–42 (2007).
 57. Borisy, A.A. *et al.* Systematic discovery of multicomponent therapeutics. *Proc. Natl. Acad. Sci. USA* **100**, 7977–7982 (2003).
 58. Kvien, T.K. *et al.* Efficacy and safety of a novel synergistic drug candidate - CRx-102 - in hand osteoarthritis. *Ann. Rheum. Dis.* **67**, 942–948 (2008).
 59. Dresser, G.K., Spence, J.D. & Bailey, D.G. Pharmacokinetic-pharmacodynamic consequences and clinical relevance of cytochrome P450 3A4 inhibition. *Clin. Pharmacokinet.* **38**, 41–57 (2000).
 60. Li, Y. *et al.* Gene expression profiling revealed novel molecular targets of docetaxel and estramustine combination treatment in prostate cancer cells. *Mol. Cancer Ther.* **4**, 389–398 (2005).
 61. Li, Y. *et al.* Gene expression profiling revealed novel mechanism of action of Taxotere and Furtulon in prostate cancer cells. *BMC Cancer* **5**, 7 (2005).
 62. Daigeler, A. *et al.* Synergistic apoptotic effects of taurolidine and TRAIL on squamous carcinoma cells of the esophagus. *Int. J. Oncol.* **32**, 1205–1220 (2008).
 63. Austin, C.P., Brady, L.S., Insel, T.R. & Collins, F.S. NIH Molecular Libraries Initiative. *Science* **306**, 1138–1139 (2004).
 64. Seiler, K.P. *et al.* ChemBank: a small-molecule screening and cheminformatics resource database. *Nucleic Acids Res.* **36**, D351–D359 (2008).
 65. Tolliday, N. *et al.* Small molecules, big players: the National Cancer Institute's Initiative for Chemical Genetics. *Cancer Res.* **66**, 8935–8942 (2006).
 66. Stark, C. *et al.* BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34**, D535–D539 (2006).
 67. Root, D.E., Flaherty, S.P., Kelley, B.P. & Stockwell, B.R. Biological mechanism profiling using an annotated compound library. *Chem. Biol.* **10**, 881–892 (2003).
 68. Taunton, J., Hassig, C.A. & Schreiber, S.L. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408–411 (1996).
 69. Miller, L.W. & Cornish, V.W. Selective chemical labeling of proteins in living cells. *Curr. Opin. Chem. Biol.* **9**, 56–61 (2005).
 70. Lefurgy, S. & Cornish, V. Finding Cinderella after the ball: a three-hybrid approach to drug target identification. *Chem. Biol.* **11**, 151–153 (2004).
 71. MacBeath, G. & Schreiber, S.L. Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760–1763 (2000).
 72. Lopez, A., Parsons, A.B., Nislow, C., Giaever, G. & Boone, C. Chemical-genetic approaches for exploring the mode of action of natural products. *Prog. Drug Res.* **66**, 237, 239–71 (2008).
 73. Campillos, M., Kuhn, M., Gavin, A.C., Jensen, L.J. & Bork, P. Drug target identification using side-effect similarity. *Science* **321**, 263–266 (2008).
 74. Melnick, J.S. *et al.* An efficient rapid system for profiling the cellular activities of molecular libraries. *Proc. Natl. Acad. Sci. USA* **103**, 3153–3158 (2006).
 75. MacDonald, M.L. *et al.* Identifying off-target effects and hidden phenotypes of drugs in human cells. *Nat. Chem. Biol.* **2**, 329–337 (2006).
 76. Schreiber, S.L. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* **287**, 1964–1969 (2000).
 77. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
 78. Kitagawa, M. *et al.* Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* **12**, 291–299 (2005).
 79. Brandner, C.J. *et al.* The ORFeome of *Staphylococcus aureus* v 1.1. *BMC Genomics* **9**, 321 (2008).
 80. Dricot, A. *et al.* Generation of the *Brucella melitensis* ORFeome version 1.1. *Genome Res.* **14**, 2201–2206 (2004).
 81. Parrish, J.R. *et al.* High-throughput cloning of *Campylobacter jejuni* ORFs by in vivo recombination in *Escherichia coli*. *J. Proteome Res.* **3**, 582–586 (2004).
 82. Murthy, T. *et al.* A full-genomic sequence-verified protein-coding gene collection for *Francisella tularensis*. *PLoS ONE* **2**, e577 (2007).
 83. Matsuyama, A. *et al.* ORFeome cloning and global analysis of protein localization in the fission yeast *Schizosaccharomyces pombe*. *Biotechnol.* **24**, 841–847 (2006).
 84. Xu, D. *et al.* Genome-wide fitness test and mechanism-of-action studies of inhibitory compounds in *Candida albicans*. *PLoS Pathog.* **3**, e92 (2007).
 85. Hoyer, L.L. & Konopka, J. *Candida* here, and *Candida* there, and *Candida* everywhere! *Future Microbiol.* **3**, 271–273 (2008).
 86. Boutos, M. *et al.* Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303**, 832–835 (2004).
 87. Reboul, J. *et al.* *C. elegans* ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat. Genet.* **34**, 35–41 (2003).
 88. Foley, E. & O'Farrell, P.H. Functional dissection of an innate immune response by a genome-wide RNAi screen. *PLoS Biol.* **2**, E203 (2004).
 89. Jones, G.M. *et al.* A systematic library for comprehensive overexpression screens in *Saccharomyces cerevisiae*. *Nat. Methods* **5**, 239–241 (2008).
 90. Brasch, M.A., Hartley, J.L. & Vidal, M. ORFeome cloning and systems biology: standardized mass production of the parts from the parts-list. *Genome Res.* **14**, 2001–2009 (2004).
 91. Tischler, J., Lehner, B. & Fraser, A.G. Evolutionary plasticity of genetic interaction networks. *Nat. Genet.* **40**, 390–391 (2008).
 92. Gardner, T.S., di Bernardo, D., Lorenz, D. & Collins, J.J. Inferring genetic networks and identifying compound mode of action via expression profiling. *Science* **301**, 102–105 (2003).
 93. Nelander, S. *et al.* Models from experiments: combinatorial drug perturbations of cancer cells. *Mol. Syst. Biol.* **4**, 216 (2008).
 94. Haggarty, S.J., Clemons, P.A. & Schreiber, S.L. Chemical genomic profiling of biological networks using graph theory and combinations of small molecule perturbations. *J. Am. Chem. Soc.* **125**, 10543–10545 (2003).
 95. St Onge, R.P. *et al.* Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nat. Genet.* **39**, 199–206 (2007).
 96. Musso, G. *et al.* The extensive and condition-dependent nature of epistasis among whole-genome duplicates in yeast. *Genome Res.* **18**, 1092–1099 (2008).
 97. Lee, M.S. *et al.* The novel combination of chlorpromazine and pentamidine exerts synergistic antiproliferative effects through dual mitotic action. *Cancer Res.* **67**, 11359–11367 (2007).
 98. Lehár, J., Krueger, A., Zimmermann, G. & Borisy, A. High-order combination effects and biological robustness. *Mol. Syst. Biol.* **4**, 215 (2008).
 99. Deutscher, D., Meilijson, I., Kupiec, M. & Ruppin, E. Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nat. Genet.* **38**, 993–998 (2006).
 100. Southan, C., Varkonyi, P. & Muresan, S. Complementarity between public and commercial databases: new opportunities in medicinal chemistry informatics. *Curr. Top. Med. Chem.* **7**, 1502–1508 (2007).
 101. Stobaugh, R.E. Chemical Abstracts Service Chemical Registry System. 11. Substance-related statistics: update and additions. *J. Chem. Inf. Comput. Sci.* **28**, 180–187 (1988).
 102. Gelperin, D.M. *et al.* Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev.* **19**, 2816–2826 (2005).
 103. Lamesch, P. *et al.* hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. *Genomics* **89**, 307–315 (2007).

