Chemistry & Biology, Vol. 11, 609–618, May, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembiol.2003.08.011

Genetic and Genomic Approaches to Identify and Study the Targets of Bioactive Small Molecules

Review

Xiaofeng S. Zheng,[1,3,*] Ting-Fung Chan,[1] and Heather H. Zhou[2]
[1]Department of Pathology and Immunology
Washington University School of Medicine
St. Louis, Missouri 63110
[2]Biotechnology and Genomics
Pfizer Global Research and Development
Chesterfield, Missouri 63017

Natural and synthetic bioactive small molecules form the backbone of modern therapeutics. These drugs primarily exert their effect by targeting cellular host or foreign proteins that are critical for the progression of disease. Therefore, a crucial step in the process of recognizing valuable new drug leads is identification of their protein targets; this is often a time consuming and difficult task. This report is intended to provide a comprehensive review of recent developments in genetic and genomic approaches to overcome the hurdle of discovering the protein targets of bioactive small molecules.

Natural and Synthetic Bioactive Small Molecules

In general, bioactive small molecules (BSMs) can be generated in two ways: these compounds are either extracted from living organisms or they are synthesized by chemical means (examples of key BSMs are shown in Figure 1). An abundance of organisms produces a diverse array of natural products, ranging from chemical compounds to small peptides, some of which have biological activity. Perhaps the most widely known class of BSM is the antibiotic; a variety of antibiotics are made by microbes, plants, and marine organisms to fight off microbial competitors or infectious agents. Rapamycin, a macrolide antibiotic that was isolated from the filamentous bacterium *S. hygroscopius* that grows in the soil of Easter Island [1], is secreted from the bacterium and induces the starvation response (characterized by G0 arrest and sporulation) in nearby fungi [2, 3], effectively allowing *S. hygroscopius* to compete with fungi for limited soil nutrients. Natural BSMs are also produced as chemical signals to control normal physiological processes such as growth. For example, eukaryotic cells synthesize the phosphosphingolipid ceramide at elevated levels, when stressed, to inhibit cell growth [4].

Both natural and synthetic bioactive compounds can have obvious medicinal value. Since the first natural product was isolated and chemically characterized, the pharmacological properties of numerous compounds have been described in animal models and in humans in clinical trials. BSMs with attractive therapeutic attributes are produced commercially and used to treat various human diseases. FK506 and cyclosporin A (CsA) are BSMs that are used in the clinic to inhibit graft rejection in kidney transplant patients. Not surprisingly, these molecules are potent inhibitors of T cell proliferation. The modulatory effects of BSMs have also been harnessed for research purposes and have been invaluable tools for probing complex biological networks and pathways leading to seminal discoveries, protein function, and protein-protein interactions. In the laboratory FK506 and CsA were instrumental in dissecting the component steps of a central signaling cascade that caused activation of T cells [5]. Both drugs bind immunophilin (an intracellular protein that binds immunosuppressive drugs) receptors, and the drug-receptor complexes inhibit calcineurin, a protein phosphatase that dephosphorylates nuclear factor of activated T cells (NF-AT), a transcription factor required for expression of interleukin-2 (IL-2) [6]. Compounds that are either too toxic or that lack the appropriate pharmacological properties for medicinal use can still be useful as research tools. Wortmannin is a metabolite derived from the fungus *T. wortmanni*. Wortmannin is renowned for the anti-inflammatory and immunosuppressive effects it exerts in experimental animals, but it has yet to advance as a treatment for humans. Nonetheless, the use of wortmannin was a critical agent for elucidating the alternative roles that the PI3K (phosphatidylinositol 3-kinase) pathway plays in aspects of cellular and organismal physiology from growth and metabolism to cell survival, as well as in disease mechanisms such as diabetes and cancer [7].

Millions of years of evolutionary selection have resulted in naturally occurring BSMs that tightly control sophisticated intra- and intercellular regulatory networks. However, many compounds with desirable properties are not amenable to large-scale production and purification due technical constraints. Consequently, combinatorial chemistry and new organic chemical synthesis protocols are generally accepted as an alternate method for producing large quantities of BSMs with diverse and complex structures [8]. With these methods, libraries composed of structurally distinct compounds can be synthesized and used to conduct meaningful target-specific or cell/organism-based phenotypic screens. The structural and chemical diversities of such compound libraries grant a high degree of assurance that a compound that can exert a desired response in a cell or on a protein will be present. One excellent example of this is the miniaturized, whole-cell immunodetection system called the cytoblot developed by Stockwell and colleagues [9]. The cytoblot was designed for high-throughput screening of a library bioactive compounds to isolate members that inhibited cell proliferation. A similar approach was used to identify monastrol, a novel inhibitor of mitotic spindle bipolarity [10]. Like natural products, the protein targets of synthetic BMSs are usually unknown. Therefore, protein target discovery is an important challenge for both the drug industry and academic chemical biology researchers.

*Correspondence: zheng@pathbox.wustl.edu
[3]Present address: Department of Pharmacology, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School, Piscataway, NJ 08854.

Sources	Small bioactive molecules	Purposes	Examples
Bacteria Fungi Plants	Natural products	Inhibit growth of competing organisms of the same habitat	Penicillin, Rapamycin, Taxol
All organisms	Messenger molecules	Intracellular signaling or cell-cell communications	Ceramide, Nitric oxide, PIPs, Acetylcholine
Organic syntheses	Synthetic molecules	Clinical drugs Chemical probes	Celebrix, Monastrol

Figure 1. An Overview of Bioactive Small Molecules

Bioactive small molecules can be obtained as natural products from bacteria, fungi, and plants, both from earth soil and sea. Living organisms also produces a wide variety of small molecules as messengers to relay signals within the cell or between cells. In recent years, synthetic chemistry also produces large numbers of bioactive small compounds that are used as therapeutics and chemical probes to study biological processes.

Unmasking the Protein Targets of Bioactive Small Molecules

To date, the vast majority of BSMs have been found to interfere with protein function, although there are some chemical compounds that bind to nucleic acids and modulate their activity. Therefore, this review will focus on the benefits and limitations of using chemical genetics and genomics to identify target proteins. Traditionally, protein targets have been isolated biochemically by using their affinity for their particular small-molecule ligands. Target proteins can be directly purified by their affinity to chemical ligands. In this approach, chemical ligands are conjugated to a solid matrix, such as Sepharose, and then used as affinity columns to withdraw target proteins from cell or tissue extracts. This approach can be very effective and has led to the discovery of important drug targets such as calcineurin [5]. Variations on this approach have also been developed, such as radioactively labeling the library compounds before screening protein-expression libraries and three-hybrid systems for affinity-based screening [11]. However, affinity purification does have shortcomings. A significant pitfall lies in the availability of chemically reactive groups that can be used for crosslinking or radiolabeling. Even when reactive groups are readily available to crosslink the compound to the solid matrix, the same groups may be needed for binding and/or to inhibit the target protein and, therefore, blocking with the solid support will render the molecule inert. Another serious roadblock that many researchers face when using affinity purification is the high background often associated with this method, especially when the affinity between BSM ligand and target protein is low. Recently, significant gains have been made in developing chemical genetic and chemical genomic approaches to characterize BSM targets. A detailed comparison of biochemical and genetic approaches is shown in Figure 2.

Comparison of biochemical and genetic approaches

Biochemical	Genetic
Strength 1. Direct identification of the target(s) 2. Suitable for all organisms Weakness 1. Limited by the availability of linker sites for conjugation to affinity matrix 2. Limited by affinity of small molecules to the target 3. Subject to artifacts generated during affinity purification	Strength 1. Targets identified are physiologically relevant 2. Can be used to identify and study pathways interfered by small molecules 3. May be used to identify and study non-protein targets Weakness 1. Targets identified may not be the direct drug targets 2. Limited to model organisms

Figure 2. A Comparison of Biochemical and Genetic Approaches in the Identification of Targets of Bioactive Small Molecules

Chemical Genetics

Genetic manipulation and analysis became a cornerstone of biological and biomedical research after comparison of several gene sequences showed that that many basic biological processes and important disease pathways are conserved among species from yeast to humans. Fundamental biological processes, including apoptosis and the cell cycle, signaling pathways, metabolism, transcription, translation, and control of cellular morphology have been described using genetics, as have disease mechanisms such as cancer and diabetes. Various methods for genetic manipulation have become standard practice for identification of key genes. The term "forward genetics" describes a screen for an individual with a desired phenotype after a population of organisms is subjected to random mutation of their genomes. A "genetic modifier screen" illuminates genes that act in a specific pathway along with a gene of interest. This is based on the principle that mutations in two genes in the same pathway would "genetically interact" with each other, causing a more deleterious or enhanced phenotype, perhaps leading to synthetic lethality if the combined mutations cause death or, alternatively, causing suppression of the phenotype associated with the initial mutation alone. If a mutation is recessive or results in loss of function for a particular protein product, the gene in question can be identified by introducing the normal genetic counterpart into the mutant strain compensating for the loss-of-function mutation, thereby producing a normal phenotype. In contrast, when the mutation has a dominant effect and affects the physiology of the organism even in the presence of a normal copy of the gene, a library of genes must be generated from the mutant strain and introduced into the wild-type organism individually to identify the mutant gene. More detailed information about chemical genetics can be found in two excellent recent reviews [12, 13].

BSMs can replicate forward genetics and genetic modifier screens (generating recessive, dominant, suppressive, enhancing, and synthetically lethal chemical mutations) by interfering with protein function, mimicking the effect of a mutation at the amino acid level. The workhorses of conventional genetics that are also tenable for screening for drug targets include the budding yeast *S. cerevisiae*, the fission yeast *S. pombe*, the

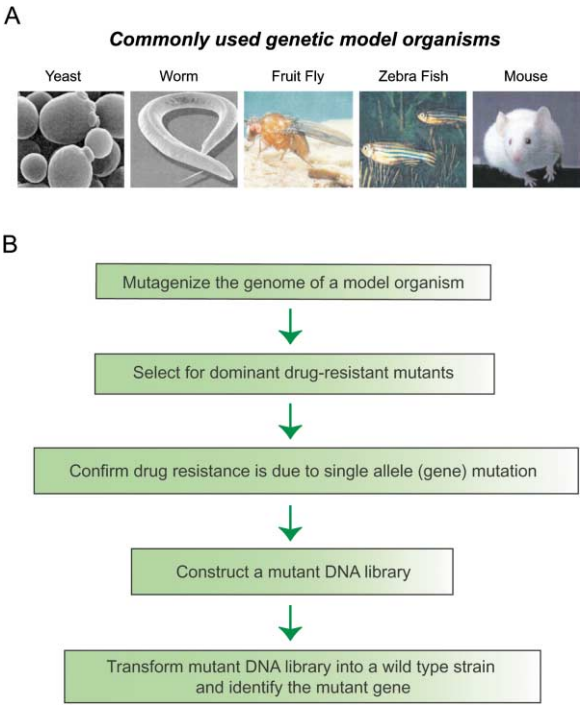

Figure 3. Genetic Model Organisms and Genetic Screens for Targets of Bioactive Small Molecules

(A) Commonly used model organisms in identification and study of drug targets. Yeast (budding yeast and fission yeast) is an excellent model system to identify and study protein targets that are conserved between fungi and animals. It is the most manipulatable organism. The nematode and fruit fly are excellent multicellular organisms. Zebrafish is increasingly used as a vertebrate model organism. All of the organisms are tenable for large-scale genetic screens. While mouse is not suitable for genetic screens, as a small mammal, it is closest to humans and is invaluable for genetic analysis of drug actions.

(B) A flow chart of a typical genetic screen for targets of bioactive small molecules. Genetic approach is a powerful way to identify and study the target of a bioactive small molecule. In a typical genetic analysis, the genome of an organism is randomly mutagenized and screened for mutations that render drug resistance or hypersensitivity. The drug-resistant and drug-sensitive genes can then be isolated by complementation. Analysis of these genes may provide important clues to the drug target.

nematode worm *C. elegans*, the fruit fly *D. melanogaster*, and the zebrafish *D. rerio* (Figure 3A).

Since small molecules mediate their effects on protein function by physical interaction, altering the small-molecule binding sites would block this interaction, resulting in a drug-resistant phenotype. The fundamental requirement for this interaction has been successfully exploited to screen and clone drug target proteins by using a forward chemical genetic approach (a general strategy is illustrated in Figure 3B) and is exemplified by the discovery of the target of rapamycin (TOR). Rapamycin is a small molecule that potently inhibits the growth of budding yeast as well as human cells. To identify the target of rapamycin, the yeast genome was screened for dominant rapamycin-resistant mutations that allowed yeast to grow in the presence of rapamycin. Mutations in two genetic alleles were discovered to permit cell growth in the presence of the drug [14]. Genomic libraries were constructed from the rapamycin-resistant strains and used to transform the wild-type yeast, leading to the identification of *TOR1* and *TOR2* [15, 16]. Tor1 and Tor2 are highly homologous, are phosphoinositol kinase-related protein kinases (PIKKs), and are conserved in mammals. Mutations at a conserved serine, Ser-1972 in *TOR1* or Ser-1975 in *TOR2*, confer dominant rapamycin resistance [14–16]. Furthermore, TOR proteins were shown to directly bind to a complex of FKBP12-rapamycin (rapamycin binds to the peptidyl-prolyl cis-trans isomerase FKBP12, resulting in a loss-of-function phenotype) [17–19]. More importantly, the serine mutations in the TOR proteins were mapped to the rapamycin binding site and were shown to disrupt binding of FKBP12-rapamycin to TOR proteins [19, 20]. Taken together, these observations unequivocally demonstrate that TOR proteins are the physiological targets of rapamycin. This example shows that drug-resistant mutations in the drug target protein are dominant because they prevent drug-target binding. Belanger and colleagues used the same principle to create a smart polymerase chain reaction (PCR)-based strategy suitable for screening for the targets of many different drugs in the pathogen *S. pneumoniae* [21]. This group generated a genomic library of *S. pneumoniae* in which random mutations were created by using error-prone PCR. The mutant library was then transformed back into the wild-type bacterium, and the bacterial transformants were grown in the presence of a normally lethal drug. The PCR products would be expressed as mutant proteins in the bugs, and an arbitrary subset of organisms would be rendered drug-resistant by the presence of certain polypeptides. Once a drug-resistant bacterium is singled out, the PCR-product conferring resistance can be determined very rapidly, significantly simplifying the target identification process. The utility of this screen was verified by using a number of known antibiotics including ciprofloxacin, fusidic acid, penicillin G, and spectinomycin, and the results confirmed some known drug targets as well as identified several potential new drug targets. This approach should be expandable to other model organisms.

Alternatively, the protein targets of BSMs may be studied by modulating the synthetic enzymes involved in BSM biosynthesis in conjunction with a genetic modifier screen. As mentioned earlier, ceramide is a eukaryotic sphingolipid with potent antiproliferative activity against tumor cells and is generated by the sphingolipid biosynthetic pathway. The small molecule ISP-1 is a newly discovered immunosuppressant isolated from the fungus *I. sinclairii*, which is a structural analog of sphingosine; ISP-1 inhibits the synthesis of ketodihydrosphingosine, a precursor for ceramide and thus blocks ceramide production [22]. Sun and colleagues isolated the gene *YPK1/SLI2* as a multicopy suppressor of ISP-1-dependent cell proliferation [22]. Ypk1 acts downstream of Pkh1, a PDK1 (3-phosphoinositide dependent)-related yeast kinase. Phytosphingosine, a normal cellular derivative of ceramide, can directly activate Pkh1 [23], thereby leading to increased phosphorylation of Ypk1 [22]. The combined use of chemical genetics and traditional genetics used in this study demonstrated that Pkh1 is a biological target of ceramide-related sphingolipids.

Confirmation of Candidate Target Proteins

Biochemical and chemical genetic approaches to identify BSM target proteins are complimentary. Biochemical affinity purification typically suffers from contamination by nonspecific proteins, and it is important to demonstrate that the purified proteins are physiologically relevant. If a small molecule inhibits a target protein's activity, it should interfere with that protein's cellular function, thus demonstrating the relevance of the target proteins identified. Trapoxin is a microbially derived cyclotetrapeptide that inhibits mammalian cell-cycle progression. mSin3 histone deacetylase (HDAC) is found to bind to trapoxin in an affinity column [24]. Treatment of cells with trapoxin leads to changes in mSin3-depedent gene expression, indicating that mSin3 HDAC is the physiologically relevant target for trapoxin [25].

The commonly used strategy of "reverse genetics" is also often used to verify purified target proteins. Reverse genetics refers to genetic manipulation of a known gene to study its various roles. If a protein were the target of a drug, certain mutations would mimic the drug effects. On the other hand, overexpression of the target gene would confer drug resistance, as in the case of camptothecin. This pentacyclic alkaloid isolated from the bark of *C. acuminata* is potently toxic to cancer cells [26]. Studies in the 1970s showed that camptothecin is a strong inhibitor of nucleic acid synthesis; however, the compound failed to inhibit purified DNA and RNA polymerases. Mammalian DNA topoisomerase (Top)I was later implicated as an intracellular target of camptothecin, since TopI could be crosslinked to the drug [27, 28]. To definitively confirm that TopI is the camptothecin target in vivo, the yeast *TOP1* gene was identified and deleted. Deletion of *TOP1* was found to confer camptothecin resistance, and overexpression of TopI causes camptothecin hypersensitivity, demonstrating that camptothecin inhibition indeed involves TopI [29, 30]. Genetically identified drug targets also require independent biochemical confirmation because mutations in other components of the drug target pathway may have additional effects on drug sensitivity. For example, Tor1 and Tor2 were initially hypothesized to act downstream of the direct rapamycin target, since the dominant rapamycin-resistant mutations occur at a conserved serine [16], and it was hypothesized that the serine mutations mimicked the phosphorylation state, resulting in a constitutively active signaling pathway causing rapamycin resistance [16]. However, as described earlier, chemical techniques demonstrated a direct interaction between the Tor proteins and the Frp1/FKBP12-rapamycin complex [19]. Most significantly, the dominant rapamycin-resistant mutations at the conserved serine (Ser-1972 in Tor1 and Ser-1975 in Tor2) were mapped to the Frp1/FKBP12-rapamycin binding site and were shown to disrupt the ability of Frp1/FKBP12-rapamycin to bind to Tor1 and Tor2 [19]. Mammalian TOR (mTOR, also called FRAP) was also isolated by FKBP12-rapamycin-affinity purification [17, 18]. Mammalian and yeast TOR proteins share a conserved 11 kDa *F*KBP12-*r*apamycin-*b*inding (FRB) domain [19, 20], which forms a four-helix-barrel structure centered with a hydrophobic pocket in which the lipophilic rapamycin is buried [31]. The conserved serine residue crucial for binding rapamycin is located at the bottom of this pocket. Substitution of this serine with a bulky amino acid such as threonine and isoleucine prevents rapamycin to FRB due to steric hindrance, resulting in dominant rapamycin resistance [19, 20, 32].

The Mechanisms of Small-Molecule Action

Genetics can also be used to discover the mechanism of action of small molecules. Brefeldin A is a fungal metabolite of *P. brefeldianum* that inhibits Golgi-mediated intracellular transport, leading to the disassembly of the Golgi complex and redistribution of Golgi proteins to the endoplasmic reticulum [33]. Brefeldin A was first shown to inhibit the guanine nucleotide exchange of ADP-ribosylation factors (Arfs), a family of small GTPases required for the COP I coat assembly, biochemically in mammalian cells [34, 35]. Genetic studies in budding yeast uncovered several brefeldin A-resistant alleles of the gene *GEA1*, a guanine nucleotide exchange factor (GEF) for Arf [36]. The brefeldin A-resistant mutations were mapped to the Arf interaction region, suggesting that brefeldin A acts by preventing the Gea1 protein product from binding to Arf. Brefeldin A binds to Arf noncompetitively, stabilizing the interaction of Arfs with its corresponding GEF and preventing the conversion from Arf-GDP to Arf-GTP [36]. The abortive affect of brefeldin A mimics a dominant-negative mutant of these GEFs; thus, brefeldin A is effectively a more efficient inhibitor than it would be by simply inhibiting Arf binding to its GEF [37]. Another good illustration of how the mechanism of action of a small molecule was uncovered is demonstrated by rapamycin inhibition of TOR proteins. The FRB domain of TOR, the binding site for FKBP12-rapamycin, is located just outside of the kinase catalytic domain, suggesting that FKBP12-rapamycin complex may directly inhibit TOR kinase activity or simply interfere with certain functions of TOR via another mechanism. In addition to the rapamycin-sensitive, essential cell-cycle activity shared with Tor1, Tor2 has a second essential but rapamycin-insensitive function in actin cytoskeleton organization [19]. Although both functions require an intact Tor2 kinase domain, only the cell cycle-related function is inhibited by rapamycin [19]. This finding indicates that rapamycin is unlikely to act as an inhibitor of the kinase catalytic activity of Tor2, which led researchers to speculate that FKBP12-rapamycin could act by disrupting the interaction between TOR and effector(s) key to cell cycle and, specifically, G1 regulation [19]. Indeed, two recent reports indicate that the interaction between mTOR and raptor, an mTOR cofactor, was disrupted by rapamycin [38, 39], providing a direct experimental support to such model.

In addition to the gene encoding the direct drug target, mutations in genes in cellular pathways or processes involving the drug target can often affect drug sensitivity, enabling the use of BSMs as probes to tease apart drug-sensitive pathways. A comprehensive understanding of the drug pathways provides a detailed molecular profile of global drug effects, including previously unrecognized benefits and potential side effects, and may lead to improved therapeutic design and clinical protocols with fewer side effects. Global cellular analysis

showed that mTOR is a nutrient sensor, regulator of diverse metabolic processes, and a component of the insulin-signaling pathway. Treating cells with rapamycin results in cellular responses that are similar to those observed under nutrient- and insulin-limitation conditions, suggesting that a potential side effect of rapamycin is type 2 diabetes. Analysis of the global effects of drugs on cells are also relevant when other biomacromolecules are the drug targets, such as DNA, RNA, and lipids. Bleomycin is an antitumor drug produced by *S. verticillus* that produces DNA double-strand breaks, leading to the selective killing of cancer cells [40]. Genetic analysis of cellular sensitivity to bleomycin and other radiomimetic drugs, together with genetic research on radiation sensitivity, reveals the presence of a cellular network of enzymes controlling DNA damage checkpoints and repair [41]. This and other similar findings are leading to improvements in cancer radiation and chemotherapeutics that are more effective in eradicating cancerous cells while providing better protection of normal cells.

Genetic Analysis of Small-Molecule Targets in Animal Models

Yeast is a unicellular eukaryotic organism with a simple genomic structure and therefore is an excellent system to study the basic cell regulatory machinery. As a "minimal" eukaryote, it has been a preferred system for the genetic analysis of drug targets and for elucidating mechanisms of drug action. However, yeast lacks a regulatory circuitry unique to multicellular organisms. For drugs whose targets are only present in multicellular animals, more complex model systems such as the fruit fly and the nematode come into play. These models can be used to study more sophisticated phenotypes that involve cell-interaction and communication between different cell types, such as hormonal regulation, organ development and function, behavior, or motor activity. Both the fruit fly and the nematode have been used to study the anesthetic halothane and related compounds [42–44]. Phenotypic and epistatic analyses suggest that volatile anesthetics interfere with the heterotrimeric G0 protein signal transduction pathway [45]. *C. elegans* has also been used to study the mechanism of action of the antidepressant fluoxetine (Prozac). Various phenotypes were observed upon fluoxetine that can be attributed to the effect of fluoxetine on serotonin reuptake, but some additional phenotypes were observed in serotonin-depleted worms, including altered neuromuscular activity [46]. A genetic screen for fluoxetine-resistant mutants led to identification of seven new genes. Analysis of these mutants revealed a family of multi-panning transmembrane proteins, which could be the fluoxetine targets resulting in the neuromuscular effect [46]. The close evolutionary relationship between the fruit fly and nematode with other insects and parasitic nematodes also make them excellent model organisms to study targets of insecticides and antithelmics [47].

The mouse is too large and has a relatively slow reproductive cycle to be useful as a model organism for screening drug targets by using genetics. Nonetheless, mice can be effective as a genetic model to verify candidate targets of BSMs. Because the mammal closely resembles humans, the mouse is an excellent genetic model for human disease, and genetically engineered mice that do not express a putative drug target gene or instead express a drug-resistant allele of the putative target gene are important for confirming drug-target interactions and investigating the effects of disease-related mutations. In addition, treatment of mutant mice with drugs may help reveal additional targets for the drugs. For instance, cyclooxygenases 1 and 2 (COX-1 and COX-2) are enzymes required for prostaglandin (PG) biosynthesis and are the targets of the widely used nonsteroidal anti-inflammatory drugs (NSAIDS). COX-1- and COX-2-deficient mice have been generated to evaluate the individual roles of each isoform in the inflammatory response [48]. COX-1 null mice showed an increase in PG concentration to about 25% above the wild-type level, whereas the COX-2 null mice showed a decrease to 75% of the wild-type level, suggesting that COX-2 is the major pathway for PG production during early stages of inflammatory responses in this model.

Drugs can induce detoxification and occasionally immunological responses. Moreover, the effectiveness of drugs can be affected by their in vivo stability and their ability to penetrate different tissues. These factors can significantly affect phenotypes of animals treated with these compounds and analysis of mutant mouse phenotypes can help predict the direct effects of small molecules or the indirect effect of target protein inhibition. Calcineurin B (CnB) is the catalytic subunit of the calcineurin phosphatase. A dominant-negative CnB (*CnB*/**) mouse was generated that showed a defect to remodel the primary vasculature to the structured pattern of the wild-type vasculature in the developing embryos [49]. Injection of CsA, an inhibitor of calcineurin, into pregnant mice between days 7.5 and 8.5, but not earlier or later, reproduced the exact vascular developmental defects seen in *CnB*/** mice [49]. The temporally selective action of CsA could not be due to degradation of the drug, failure of placental transfer, or embryonic metabolism of the drugs, since similar levels of CsA were achieved during the critical period and after it. The observation that CsA administration mimics the phenotype of the *CnB*/** mutants indicates that in early mammalian development, CsA is a highly specific inhibitor of calcineurin function and is unlikely to have other developmentally critical targets.

Genomic Drug-Sensitivity Screens

The complete sequence of the human genome and many other genomes has been reported in the last few years. Knowledge of the sequence of entire genomes has fundamentally changed the landscape of biological and biomedical research. Genomic and proteomic tools such as large-scale gene knockouts, DNA, and protein microarrays provide ways to systematically study gene and protein functions. Likewise, these new research tools are being rapidly integrated into discovery and study of drug target proteins. After the completion of genomic sequencing, efforts are being made to generate systematic deletion mutants of model organisms such as budding yeast, worms, fruit flies, and mice by using

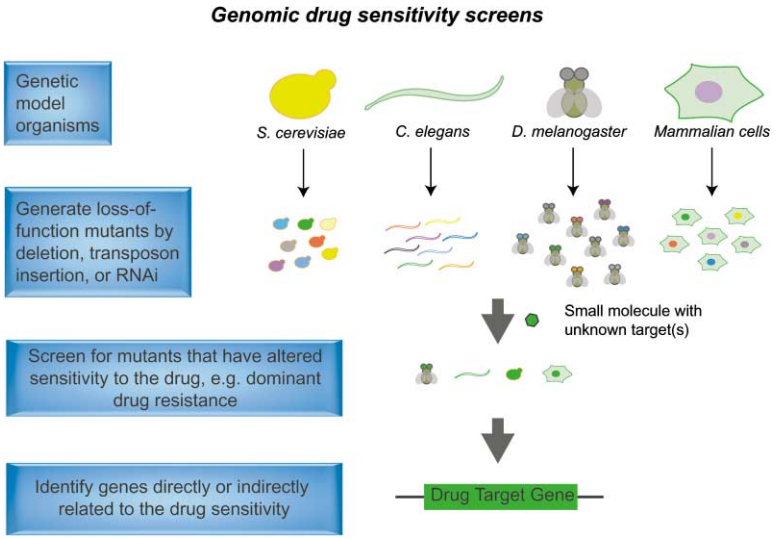

Figure 4. Genomic Drug-Sensitivity Screens

Genomic deletion mutants using homologous recombination or RNAi can be used directly in drug-sensitivity screens. Mutation of drug target genes or genes acting in the biological pathways interfered with by bioactive small molecules is known to alter drug sensitivity.

strategies from homologous recombination and transposon insertion to RNAi [50–54]. The *Saccharomyces* Genome Deletion Project generated the first complete sets of deletion mutants of an organism [53]. Deletion of target gene(s) confers drug resistance or hypersensitivity, depending on the nature of the drug action, and therefore the deletion mutants provide an exceptional opportunity for systematically identifying and studying drug targets [55, 56] (Figure 4). Deletion mutants may be arrayed and assayed for drug sensitivity on drug-containing agar plates or assayed in liquid cultures in multiwell plates [57]. High-throughput strategies for such screens have recently been developed, such as cell microarrays that were recently been created to accommodate large numbers of samples in small surface areas [58]. In addition, the yeast deletion mutants are each tagged with a unique 20-base oligonucleotide sequence [53, 59]. A high-density oligonucleotide microarray recognizing the tag sequences has been generated to analyze the relative abundance of individual deletion strains in a cocktail of mutants in the presence of drugs. Giaever and colleagues conducted a proof-of-concept study by using 233 heterozygous diploid strains to examine tunicamycin sensitivity through a haploinsufficiency approach [53]. They identified Alg7, an Asn-linked glycosyltransferase in yeast and a known tunicamycin target. On the same note, our laboratory recently screened the entire yeast deletion-mutant collections for rapamycin sensitivity and found that both the *tor1*Δ mutation in haploid yeast and the *tor2*Δ mutation in heterozygous diploid yeast rendered the organism hypersensitive to rapamycin [57]. Pioneering studies like those described above established the feasibility of using genomic deletion mutants in drug target discovery.

Mutations in other genes in drug-sensitive pathways normally affect the host's response to the drug, although to a lesser degree than mutations in drug target genes. Therefore, genomic deletion mutants are powerful for systematically identifying pathways influenced by bioactive small molecules. Complete knowledge of all the proteins sensitive to particular drug in an organism may allow assembly of genetic pathways involving the drug target, thereby allowing assessment of the global effects of drugs on cells and organisms [55, 56]. Additionally, drug-sensitive gene products could be determining factors dictating the responses of certain diseases to specific therapies. Establishing a profile of drug sensitivity genes, perhaps in cancer patients, could allow better prognosis of the outcome of chemotherapy.

DNA Microarrays in Small-Molecule Target Research

High-density oligonucleotide or complementary DNA (cDNA) microarrays, collectively called DNA microarrays, are used to simultaneously monitor expression of genes in a genome [60]. DNA microarrays have been extensively used to generate global gene expression profiles under diverse cellular conditions, stress, and stimulation with various extracellular stimuli. Changes in global gene-expression profiles can provide important clues about cellular processes and functions of relevant genes. DNA microarrays have been shown to be useful in the identification, validation, and study of small-molecule targets. The immunosuppressive drugs FK506 and cyclosporin A (CsA) inhibit calcineurin, a protein phosphatase critical for T cell signaling to regulate IL-2 expression [5]. DNA microarray analysis showed that treatment of a wild-type yeast strain with FK506 and CsA alters gene expression profiles that phenocopy those caused by the null mutations of *CNA1* and *CNA2*, the calcineurin genes [61]. When FK506 was used at higher concentration, a number of other genes were further induced independent of calcineurin [61]. These genes are known to be regulated by the transcription factor Gcn4, suggesting that Gcn4 is a secondary target of FK506. One of the difficulties associated with the DNA microarray approach is the scattered large amount of unorganized data. To automate the data analytical process, Hughes and colleagues constructed a database called a "compendium" of 300 expression profiles corresponding to diverse mutations and chemical treatments of budding yeast [62] (Figure 5). By grouping expression profiles with similar patterns, they identified one gene, *YER044C*, whose null mutation exhibits an expression

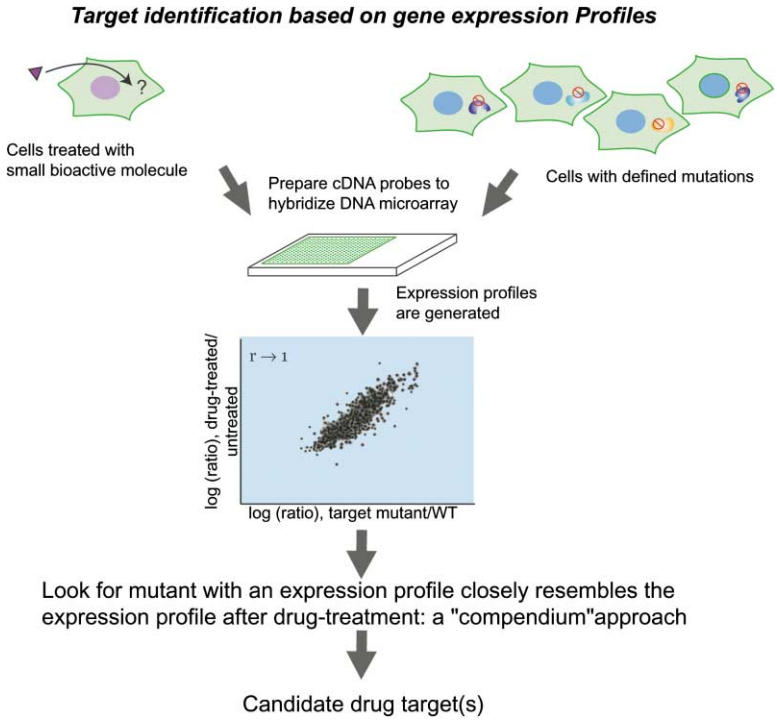

Figure 5. Gene Expression Profile-Based Identification of Targets of Small Molecules

In a so-called compendium approach, A gene expression profile of drug-treated cells is compared to expression profiles of mutant cells. A match of the expression profile of drug-treated cells with one by gene mutation would lead to a candidate drug target.

profile similar to mutations in *ERG2*, *ERG3*, and *ERG11*, components of the ergosterol pathway. They therefore named *YER044C* as *ERG28* and demonstrated by gas chromatography that *erg28* mutant cells have an unusual ergosterol content in comparison to the wild-type cells, indicating that the Erg28 protein is a new component in ergosterol metabolism. In a separate experiment, these authors discovered that the expression profile of *erg2*Δ mutant was similar to that of the wild-type strain treated with dyclonine, a topical anesthetic whose target was unknown. Gas chromatography confirmed that dyclonine treatment causes accumulation of fecosterol in wild-type cells, indicating dyclonine inhibits Erg2, a sterol C-8 isomerase. This appropriately named "relevance network" approach was further developed to identify genes linked to anticancer drug susceptibility [63]. Butte and colleagues generated expression profiles of 7245 genes in the NCI60 cells, a set of 60 different human cancer cell lines compiled by the National Cancer Institute (NCI). In a separate study, a database of gene expression profiles was generated for the same cancer cell lines treated with 5084 anticancer agents at half-growth inhibitory concentration (GI50). By comparing the two databases, a network of genes was identified. Changes in their expression levels may be linked to cellular responses to a particular anticancer drug at GI50. *LCP1* is one of the genes that showed significant changes in response to the anticancer agent NSC624044. *LCP1* encodes L-plastin, an actin binding protein involved in leukocyte adhesion. Cells overexpressing *LCP1* are found to have a higher GI50 for NSC624044. Therefore, potential target proteins may be identified by comparing the expression profiles of defined mutants with that of cells treated with bioactive small molecules.

Proteomic Approaches to Reveal Small Molecule Targets

Recent advances in the components of the proteome of different organisms and technologies to detect the identity of proteins have greatly facilitated discovery of BSM protein targets. Proteins purified by drug-affinity column can be readily identified by mass spectrometry (MS) with very high sensitivity. Detailed information on MS can be found in a recent review [64]. Furthermore, protein microarrays or the protein chip has been fabricated by using the same concept as DNA microarray. Like DNA microarray, distinct proteins in small quantity are arrayed on microchips to generate protein microarrays [65, 66]. Protein chips have been successfully used to identify protein kinase substrates [66] and antigens [67], and they are also suitable for identifying drug targets since they can be directly used to assay for the binding of bioactive small molecules to a proteome array (Figure 6). In a typical binding study, small chemicals are labeled with color matrices and used to measure affinity to thousands of proteins on a microchip [65]. After stringent washing, the proteins significantly labeled by the chemicals will be identified by their positions on the chip. Theoretically, protein chips are perhaps the most direct approach to identify drug targets [68]; however, to provide an accurate representation of the cellular milieu, microarrays require sufficient representation of purified recombinant proteins on a proteomic scale. A major difficulty for this technology is that many eukaryotic proteins, especially proteins with large molecular weight, are difficult to produce in *E. coli*. On the other hand, purification of native proteins or recombinant proteins from eukaryotic systems may be complicated by the presence of copurified proteins in a complex or of contaminating proteins. Therefore, while

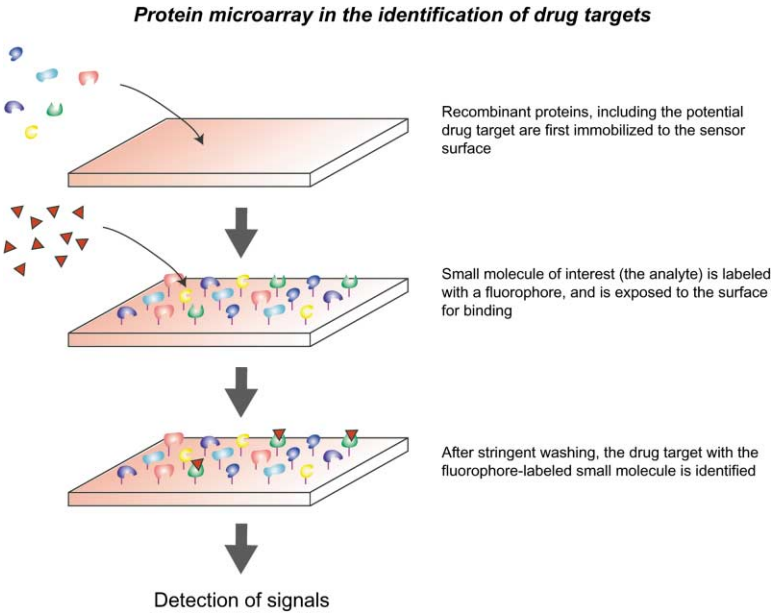

Figure 6. Protein Microarray in the Identification of Drug Targets

Recombinant proteins are arrayed onto small surface areas to generate protein microarray. Fluorescence-labeled bioactive small molecules are used to identify their target proteins by their affinity.

protein microarrays hold perhaps the greatest promises, their practical use for drug target identification still awaits technological improvements.

Conclusions

Bioactive small molecules have considerable value for therapeutic intervention and chemical biological research. Genetic and genomic approaches to identify the targets of these compounds are most effective when used in conjunction with biochemical approaches. Candidate drug targets can be confirmed by their affinity to the chemical compounds. On the other hand, proteins found to bind certain drugs depend on chemical genetic and genomic analyses to demonstrate their physiological relevance of the drug's effects. Target identification is an important first step toward generating a molecular understanding of the mechanism of drug action, the therapeutic benefits, and side effects (Figure 7). This also often leads to groundbreaking insights into biological processes and pathways targeted by bioactive small molecules, through which we can learn a great deal about basic cell regulation, physiology, and disease mechanisms. Such new knowledge can then be used to discover new targets for developing better medicine.

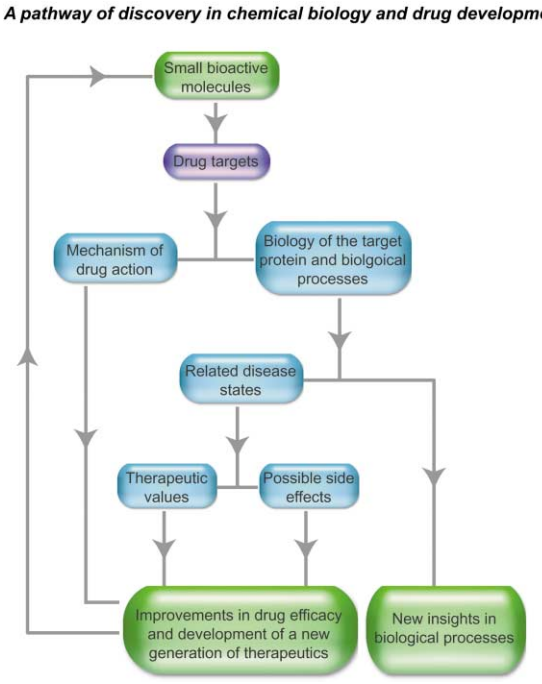

Figure 7. A Genetic Road Map to Drug Targets

A flowchart shows the importance of drug target discovery in the study of drug mechanism of action, therapeutic outcomes, new drug development, and basic biological processes.

Acknowledgments

The authors would like to thank J. Carvalho for reading the manuscript. Work in the authors' laboratory is supported by grants from the National Institutes of Health (RO1CA77668, RO1CA099004, and RO1GM62817) and the Department of Defense (X.S.Z.).

References

1. Sehgal, S. (2003). Sirolimus: its discovery, biological properties, and mechanism of action. Transplant. Proc. *35*, S7–S14.
2. Barbet, N.C., Schneider, U., Helliwell, S.B., Stansfield, I., Tuite, M.F., and Hall, M.N. (1996). TOR controls translation initiation and early G1 progression in yeast. Mol. Biol. Cell *7*, 25–42.
3. Zheng, X.F., and Schreiber, S.L. (1997). Target of rapamycin proteins and their kinase activities are required for meiosis. Proc. Natl. Acad. Sci. USA *94*, 3070–3075.
4. Hannun, Y.A., and Obeid, L.M. (2002). The ceramide-centric universe of lipid mediated cell regulation: stress encounters of the lipid kind. J. Biol. Chem. *277*, 25847–25850.
5. Liu, J., Farmer, J.D., Lane, W.S., Friedman, J., Weissman, I., and Schreiber, S.L. (1991). Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. Cell *66*, 807–815.

6. Crabtree, G.R. (2001). Calcium, calcineurin, and the control of transcription. J. Biol. Chem. *276*, 2313–2316.
7. Fruman, D.A., Meyers, R.E., and Cantley, L.C. (1998). Phosphoinositide kinases. Annu. Rev. Biochem. *67*, 481–507.
8. Schreiber, S. (2000). Target-oriented and diversity-oriented organic synthesis in drug discovery. Science *287*, 1964–1969.
9. Stockwell, B.R., Haggarty, S.J., and Schreiber, S.L. (1999). High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. Chem. Biol. *6*, 71–83.
10. Mayer, T.U., Kapoor, T.M., Haggarty, S.J., King, R.W., Schreiber, S.L., and Mitchison, T.J. (1999). Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. Science *286*, 971–974.
11. Licitra, E.J., and Liu, J.O. (1996). A three-hybrid system for detecting small ligand-protein receptor interactions. Proc. Natl. Acad. Sci. USA *93*, 12817–12821.
12. Crews, C.M., and Splittgerber, U. (1999). Chemical genetics: exploring and controlling cellular processes with chemical probes. Trends Biochem. Sci. *24*, 317–320.
13. Stockwell, B.R. (2000). Chemical genetics: ligand-based discovery of gene function. Nat. Rev. Genet. *1*, 116–125.
14. Heitman, J., Movva, N.R., and Hall, M.N. (1991). Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. Science *253*, 905–909.
15. Cafferkey, R., Young, P.R., McLaughlin, M.M., Bergsma, D.J., Koltin, Y., Sathe, G.M., Faucette, L., Eng, W.K., Johnson, R.K., and Livi, G.P. (1993). Dominant missense mutations in a novel yeast protein related to mammalian phosphatidylinositol 3-kinase and VPS34 abrogate rapamycin cytotoxicity. Mol. Cell. Biol. *13*, 6012–6023.
16. Kunz, J., Henriquez, R., Schneider, U., Deuter-Reinhard, M., Movva, N.R., and Hall, M.N. (1993). Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase homolog required for G1 progression. Cell *73*, 585–596.
17. Brown, E.J., Albers, M.W., Shin, T.B., Ichikawa, K., Keith, C.T., Lane, W.S., and Schreiber, S.L. (1994). A mammalian protein targeted by G1-arresting rapamycin-receptor complex. Nature *369*, 756–758.
18. Sabatini, D.M., Erdjument-Bromage, H., Lui, M., Tempst, P., and Snyder, S.H. (1994). RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin dependent fashion and is homologous to yeast TORs. Cell *78*, 35–43.
19. Zheng, X.F., Florentino, D., Chen, J., Crabtree, G.R., and Schreiber, S.L. (1995). TOR kinase domains are required for two distinct functions, only one of which is inhibited by rapamycin. Cell *82*, 121–130.
20. Chen, J., Zheng, X.F., Brown, E.J., and Schreiber, S.L. (1995). Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. Proc. Natl. Acad. Sci. USA *92*, 4947–4951.
21. Belanger, A.E., Lai, A., Brackman, M.A., and LeBlanc, D.J. (2002). PCR-based ordered genomic libraries: a new approach to drug target identification for streptococcus pneumoniae. Antimicrob. Agents Chemother. *46*, 2507–2512.
22. Sun, Y., Taniguchi, R., Tanoue, D., Yamaji, T., Takematsu, H., Mori, K., Fujita, T., Kawasaki, T., and Kozutsumi, Y. (2000). Sli2 (Ypk1), a homologue of mammalian protein kinase SGK, is a downstream kinase in the sphingolipid-mediated signaling pathway of yeast. Mol. Cell. Biol. *20*, 4411–4419.
23. Friant, S., Lombardi, R., Schmelzle, T., Hall, M.N., and Riezman, H. (2001). Sphingoid base signaling via Pkh kinases is required for endocytosis in yeast. EMBO J. *20*, 6783–6792.
24. Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science *272*, 408–411.
25. Hassig, C.A., Fleischer, T.C., Billin, A.N., Schreiber, S.L., and Ayer, D.E. (1997). Histone deacetylase activity is required for full transcriptional repression by mSin3A. Cell *89*, 341–347.
26. Wall, M. (1998). Camptothecin and taxol: discovery to clinic. Med. Res. Rev. *18*, 299–314.
27. Hsiang, Y.H., Hertzberg, R., Hecht, S., and Liu, L.F. (1985). Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I. J. Biol. Chem. *260*, 14873–14878.
28. Hsiang, Y.H., and Liu, L.F. (1988). Identification of mammalian DNA topoisomerase I as an intracellular target of the anticancer drug camptothecin. Cancer Res. *48*, 1722–1726.
29. Nitiss, J., and Wang, J.C. (1988). DNA topoisomerase-targeting antitumor drugs can be studied in yeast. Proc. Natl. Acad. Sci. USA *85*, 7501–7505.
30. Knab, A.M., Fertala, J., and Bjornsti, M.A. (1993). Mechanisms of camptothecin resistance in yeast DNA topoisomerase I mutants. J. Biol. Chem. *268*, 22322–22330.
31. Choi, J., Chen, J., Schreiber, S.L., and Clardy, J. (1996). Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. Science *273*, 239–242.
32. Brown, E.J., Beal, P.A., Keith, C.T., Chen, J., Shin, T.B., and Schreiber, S.L. (1995). Control of p70 s6 kinase by kinase activity of FRAP in vivo. Nature *377*, 441–446.
33. Pelham, H. (1991). Multiple targets for brefeldin A. Cell *67*, 449–451.
34. Donaldson, J.G., Finazzi, D., and Klausner, R.D. (1992). Brefeldin A inhibits Golgi membrane-catalysed exchange of guanine nucleotide onto ARF protein. Nature *360*, 350–352.
35. Helms, J.B., and Rothman, J.E. (1992). Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF. Nature *360*, 352–354.
36. Peyroche, A., Antonny, B., Robineau, S., Acker, J., Cherfils, J., and Jackson, C.L. (1999). Brefeldin A acts to stabilize an abortive ARF-GDP-Sec7 domain protein complex: involvement of specific residues of the Sec7 domain. Mol. Cell *3*, 275–285.
37. Chardin, P., and McCormick, F. (1999). Brefeldin A: the advantage of being uncompetitive. Cell *97*, 153–155.
38. Kim, D., Sarbassov, D.D., Ali, S., King, J., Latek, R., Erdjument-Bromage, H., Tempst, P., and Sabatini, D. (2002). mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. Cell *110*, 163–175.
39. Hara, K., Maruki, Y., Long, X., Yoshino, K., Oshiro, N., Hidayat, S., Tokunaga, C., Avruch, J., and Yonezawa, K. (2002). Raptor, a binding partner of target of rapamycin (TOR), mediates TOR action. Cell *110*, 177–189.
40. Hecht, S. (2000). Bleomycin: new perspectives on the mechanism of action. J. Nat. Prod. *63*, 158–168.
41. Hartwell, L., Weinert, T., Kadyk, L., and Garvik, B. (1994). Cell cycle checkpoints, genomic integrity, and cancer. Cold Spring Harb. Symp. Quant. Biol. *59*, 259–263.
42. Campbell, D.B., and Nash, H.A. (1994). Use of Drosophila mutants to distinguish among volatile general anesthetics. Proc. Natl. Acad. Sci. USA *91*, 2135–2139.
43. Krishnan, K.S., and Nash, H.A. (1990). A genetic study of the anesthetic response: mutants of Drosophila melanogaster altered in sensitivity to halothane. Proc. Natl. Acad. Sci. USA *87*, 8632–8636.
44. van-Swinderen, B., Shook, D.R., Ebert, R.H., Cherkasova, V.A., Johnson, T.E., Shmookler-Reis, R.J., and Crowder, C.M. (1997). Quantitative trait loci controlling halothane sensitivity in Caenorhabditis elegans. Proc. Natl. Acad. Sci. USA *94*, 8232–8237.
45. van Swinderen, B., Metz, L.B., Shebester, L.D., Mendel, J.E., Sternberg, P.W., and Crowder, C.M. (2001). Goalpha regulates volatile anesthetic action in Caenorhabditis elegans. Genetics *158*, 643–655.
46. Choy, R.K., and Thomas, J.H. (1999). Fluoxetine-resistant mutants in C. elegans define a novel family of transmembrane proteins. Mol. Cell *4*, 143–152.
47. Kenerson, H.L., Aicher, L.D., True, L.D., and Yeung, R.S. (2002). Activated mammalian target of rapamycin pathway in the pathogenesis of tuberous sclerosis complex renal tumors. Cancer Res. *62*, 5645–5650.
48. Langenbach, R., Loftin, C., Lee, C., and Tiano, H. (1999). Cyclooxygenase knockout mice: models for elucidating isoform-specific functions. Biochem. Pharmacol. *58*, 1237–1246.
49. Graef, I.A., Chen, F., Chen, L., Kuo, A., and Crabtree, G.R. (2001). Signals transduced by Ca(2+)/calcineurin and NFATc3/c4 pattern the developing vasculature. Cell *105*, 863–875.
50. Dong, L.Q., Zhang, R.B., Langlais, P., He, H., Clark, M., Zhu, L., and Liu, F. (1999). Primary structure, tissue distribution, and

expression of mouse phosphoinositide-dependent protein kinase-1, a protein kinase that phosphorylates and activates protein kinase Czeta. J. Biol. Chem. *274*, 8117–8122.

51. Bargmann, C.I. (2001). High-throughput reverse genetics: RNAi screens in Caenorhabditis elegans. Genome Biol *2*, REVIEWS1005.1–1005.3.
52. Spradling, A.C., Stern, D., Beaton, A., Rhem, E.J., Laverty, T., Mozden, N., Misra, S., and Rubin, G.M. (1999). The Berkeley Drosophila genome project gene disruption project: single P-element insertions mutating 25% of vital Drosophila genes. Genetics *153*, 135–177.
53. Giaever, G., Shoemaker, D.D., Jones, T.W., Liang, H., Winzeler, E.A., Astromoff, A., and Davis, R.W. (1999). Genomic profiling of drug sensitivities via induced haploinsufficiency. Nat. Genet. *21*, 278–283.
54. Coelho, P.S., Kumar, A., and Snyder, M. (2000). Genome-wide mutant collections: toolboxes for functional genomics. Curr. Opin. Microbiol. *3*, 309–315.
55. Zheng, X.F., and Chan, T.-F. (2001). Chemical genomics in the global study of protein functions. Drug Discov. Today *7*, 197–205.
56. Zheng, X.F., and Chan, T.-F. (2001). Chemical genomics: a systematic approach in biological research and drug discovery. Curr. Issues Mol. Biol. *4*, 33–43.
57. Chan, T.F., Carvalho, J., Riles, L., and Zheng, X.F.S. (2000). A chemical genomics approach toward understanding the global functions of TOR. Proc. Natl. Acad. Sci. USA *97*, 13227–13232.
58. Xu, C.W. (2002). High-density cell microarrays for parallel functional determinations. Genome Res. *12*, 482–486.
59. Shoemaker, D.D., Lashkari, D.A., Morris, D., Mittmann, M., and Davis, R.W. (1996). Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. Nat. Genet. *14*, 450–456.
60. Lockhart, D.J., and Winzeler, E.A. (2000). Genomics, gene expression and DNA arrays. Nature *405*, 827–836.
61. Marton, M.J., DeRisi, J.L., Bennett, H.A., Iyer, V.R., Meyer, M.R., Roberts, C.J., Stoughton, R., Burchard, J., Slade, D., Dai, H., et al. (1998). Drug target validation and identification of secondary drug target effects using DNA microarrays. Nat. Med. *4*, 1293–1301.
62. Bernstein, B.E., Tong, J.K., and Schreiber, S.L. (2000). Genomewide studies of histone deacetylase function in yeast. Proc. Natl. Acad. Sci. USA *97*, 13708–13713.
63. Butte, A.J., Tamayo, P., Slonim, D., Golub, T.R., and Kohane, I.S. (2000). Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks. Proc. Natl. Acad. Sci. USA *97*, 12182–12186.
64. Graves, P.R., and Haystead, T.A.J. (2002). Molecular biologist's guide to proteomics. Microbiol. Mol. Biol. Rev. *66*, 39–63.
65. MacBeath, G., and Schreiber, S.L. (2000). Printing proteins as microarrays for high-throughput function determination. Science *289*, 1760–1763.
66. Zhu, H., Klemic, J.F., Chang, S., Bertone, P., Casamayor, A., Klemic, K.G., Smith, D., Gerstein, M., Reed, M.A., and Snyder, M. (2000). Analysis of yeast protein kinases using protein chips. Nat. Genet. *26*, 283–289.
67. Lueking, A., Horn, M., Eickhoff, H., Bussow, K., Lehrach, H., and Walter, G. (1999). Protein microarrays for gene expression and antibody screening. Anal. Biochem. *270*, 103–111.
68. Weinberger, S.R., Morris, T.S., and Pawlak, M. (2000). Recent trends in protein biochip technology. Pharmacogenomics *1*, 395–416.