Adding Precision Tools to the Plant Biologists' Toolbox with Chemical Genomics

Natasha Raikhel* and Michael Pirrung

Botany and Plant Sciences Department (N.R.), Chemistry Department (M.P.), and Center for Plant Cell Biology (N.R., M.P.), University of California, Riverside, California 92521

The scope of our knowledge regarding plant genetic mechanisms and gene functions is rapidly evolving and expanding. This is primarily due to the availability of genetic and genomic methods for investigating model species, such as Arabidopsis (*Arabidopsis thaliana*). To fuel future advances, current gene-based methods must be complemented by innovative interdisciplinary approaches that are broadly applicable to dynamic and complex biological processes and functional genomics analyses. We believe that the use of diverse chemicals to interrogate molecular processes provides a novel avenue for the rapid and effective dissection of biological mechanisms and gene networks in ways not feasible with mutation-based approaches. By facilitating the identification of new pathways and networks, this powerful technology, called chemical genomics, overcomes important gaps in ongoing functional genomics efforts in plants and allows for the eventual development of a framework for predictive modeling.

The chemical genomics approach uses small molecules to modify or disrupt the functions of specific genes/proteins (Stockwell, 2000; Dobson, 2004; Lipinski and Hopkins, 2004), in contrast to classic genetics, in which mutations disrupt gene function. The underlying concept is that the functions of most proteins can be altered by the binding of a chemical, which can be found by screening large libraries for compounds that specifically affect a measurable process. There are four major aspects to chemical genomics: (1) library assembly/synthesis, or the creation of chemically diverse libraries of compounds; (2) screening, or the identification of compounds that affect a biological process of interest; (3) target identification, or the discovery of the protein targets of active compounds; and (4) target function and network discovery, or the use of the compounds to understand biological processes. It is usually necessary to screen a large number of compounds to find one or a few of sufficient specificity and efficacy to be useful, analogous to genetically screening for mutations causing a specific phenotype. However, the chemical genomics approach can address loss-of-function lethality and gene redundancy and allow instantaneous, reversible, tunable, and conditional control of a phenotype, providing many advantages over traditional genetic approaches. Well-characterized bioactive chemicals and their targets identified in Arabidopsis can be used in non-model species to improve agronomic traits and increase crop value.

Bioactive chemicals have a long history of helping plant physiologists unravel mechanisms, including those involving inhibitors of GA biosynthesis, inhibitors of ethylene action, inhibitors of auxin transport, cytoskeleton-disrupting drugs, and inhibitors of GDP-GTP exchange proteins, just to name a few. However, this approach has also met with strong criticism due to the complexities associated with understanding the action mode of compounds at the molecular level. This is one reason why drug companies must advertise the side effects of the drugs they sell. What has motivated biologists to revisit their interest in small molecules?

While a little more than 10 million pure compounds are known in chemical literature, the potential chemical diversity (defined as the number of unique chemical structures) of compounds composed of carbon, hydrogen, nitrogen, oxygen, sulfur, phosphorous, and the halogens (the organic chemist's periodic table) of molecular weight $<$1,000 likely exceeds 10^{60}. The compounds that have thus far been tested for effects on plants are therefore only a minute fraction of the structural possibilities. The development of combinatorial and automated techniques for synthesizing novel compounds brought forth significant enhancement in the productivity of chemists and makes the likelihood of synthesizing molecular libraries that are representative of "chemical space" much greater. These advances in technology allow a systematic analysis of these chemicals. A more systematic approach means that we discover chemicals that specifically disrupt a process or the function of a protein. Once these chemicals are identified, we can combine their use with genetic screens to identify genes involved in the same process. The use of unbiased libraries of diverse small molecules will allow plant biologists to discover numerous new bioactive molecules valuable for studying the function of uncharacterized plant genes. Importantly, when combined with Arabidopsis functional genomics, chemical genomics is powerful for the effective and efficient analysis of regulatory networks underlying a specific process.

Chemical genomics technologies have been used by industry for a long time. The only academic institutions devoted to this approach with a focus on

* Corresponding author; e-mail nraikhel@ucr.edu; fax 951–827–2155.

www.plantphysiol.org/cgi/doi/10.1104/pp.104.900155.

mammalian cells and microorganisms are the Broad Chemical Biology Platform (former Institute for Chemical Biology) of Harvard University and the National Institutes of Health Chemical Genomics Center (Dobson, 2004). Chemical genetics approaches have been used to dissect pathways in single-celled systems, including bacteria, yeast, and mammalian cell cultures and in *Caenorhabditis elegans* and zebrafish. Several studies have shown that identified chemicals can exhibit extremely high specificity toward target proteins. The major challenge in mammalian systems is the identification of target proteins because of the lack of an efficient molecular genetic/genomic approach (Stockwell, 2000). This problem has been acknowledged by recent approaches to develop knockout collections in mice (Austin et al., 2004). Several recent reports in the field of plant biology, however, have demonstrated the feasibility of chemical genomics in Arabidopsis (Armstrong et al., 2004; Zouhar et al., 2004; Surpin et al., 2005), but the approach has not yet been widely accepted by plant biologists.

We argue that chemical genomics has several major advantages when integrated with Arabidopsis genomics/proteomics tools, including the capabilities to: (1) effectively dissect a complex gene network affected by chemicals of interest; (2) rapidly identify their gene targets; and (3) efficiently investigate their mode of action. The reference plant Arabidopsis with its small body size and high seed yield is ideal for chemical screens at various biological levels ranging from subcellular to the whole organism. Arabidopsis offers many unique genetic/genomic tools, including a completely sequenced genome (Arabidopsis Genome Initiative, 2000), whole-genome microarrays, a large collection of knockout and activation-tagged mutations, a rich array of mutants, and the ease of cloning a gene by map-based cloning aided by DNA microarrays. Arabidopsis was invaluable for the identification of receptors for brassinosteroids (Wang et al., 2001) and ethylene (Schaller and Bleecker, 1995), just to name a couple.

Critical to the chemical genomics approach is the ability to screen for mutants resistant or hypersensitive to bioactive chemicals and to identify responsible genes using genomics-based methods. Mutations conferring resistance/hypersensitivity can either affect the target of the chemical, or upstream and downstream components of its target. Once responsible genes are cloned, cognate target proteins may be identified by further studies, such as chemical binding assays. Importantly, mutations in a component upstream/downstream of the target protein will not only give insights into the function of target genes but also provide a new high-throughput method for identifying new genes in the corresponding network. In other words, chemical genomics will enhance the power of existing genomic tools including the large collections of T-DNA insertion lines, from which we can recover mutants with altered chemical sensitivity that otherwise may not exhibit a phenotype (Blackwell and Zhao, 2003). Thus, the integration of chemical genomics with Arabidopsis genomics provides a strategic advantage and a powerful high-throughput approach for gene network discovery, target identification, and mode-of-action studies.

Of course, it is necessary to use chemical genomics with caution, and the tool has its own limitations. For example, the specificity of the selected compounds needs to be addressed, and possible uptake and metabolism of the compounds may affect the discovery and analysis of bioactive compounds. Consideration of known mechanisms of metabolism suggests strategies for the design of metabolism-resistant small molecule libraries. For example, fluorinated or cyclopropyl compounds that are resistant to cytochrome P450 oxidation are well known. Such oxidation-resistant groupings can be used in the construction of chemical libraries to prevent metabolism. This will help assure that mutants that are resistant to a ''hit'' chemical have modified targets (or at least modified proteins in the same pathway). As with any new technology, chemical genomics has its pros and cons. However, we believe that chemical genomics offers a strategy to extend the applicability of functional genomics in plants by addressing issues of overlapping gene function in gene families, lethal loci, and control of dosage- and tissue/development-specific applications.

LITERATURE CITED

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408:** 796–815

Armstrong JI, Yuan S, Dale JM, Tanner VN, Theologis A (2004) Identification of inhibitors of auxin transcriptional activation by means of chemical genetics in Arabidopsis. Proc Natl Acad Sci USA **101:** 14978–14983

Austin CP, Battey JF, Bradley A, Bucan M, Capecchi M, Collins FS, Dove WF, Duyk G, Dymecki S, Eppig JT, et al (2004) The knockout mouse project. Nat Genet **36:** 921–924

Blackwell HE, Zhao Y (2003) Chemical genetic approaches to plant biology. Plant Physiol **133:** 448–455

Dobson CM (2004) Chemical space and biology. Nature **432:** 824–828

Lipinski C, Hopkins A (2004) Navigating chemical space for biology and medicine. Nature **432:** 855–861

Schaller GE, Bleecker AB (1995) Ethylene-binding sites generated in yeast expressing the Arabidopsis ETR1 gene. Science **270:** 1809–1811

Stockwell BR (2000) Chemical genetics: ligand-based discovery of gene function. Nat Rev Genet **1:** 116–125

Surpin M, Rojas-Pierce M, Carter C, Hicks GR, Vasquez J, Raikhel NV (2005) The power of chemical genomics to study the link between endomembrane system components and the gravitropic response. Proc Natl Acad Sci USA **102:** 4902–4907

Wang ZY, Seto H, Fujioka S, Yoshida S, Chory J (2001) BRI1 is a critical component of a plasma-membrane receptor for plant steroids. Nature **410:** 380–383

Zouhar J, Hicks GR, Raikhel NV (2004) Sorting inhibitors (Sortins): chemical compounds to study vacuolar sorting in Arabidopsis. Proc Natl Acad Sci USA **101:** 9497–9501