CONCEPT OF THE IONOME

Living systems are supported and sustained by the genome through the action of the transcriptome, proteome, metabolome, and ionome, the four basic biochemical pillars of functional genomics. These pillars represent the sum of all the expressed genes, proteins, metabolites, and elements within an organism. The dynamic response and interaction of these biochemical "omes" defines how a living system functions, and its study, systems biology, is now one of the biggest challenges in the life sciences. Studies on the functional connections between the genome and the transcriptome (Marti*vanou and Hampp, 2003; Becher et al., 2004; Leonhardt et al., 2004), proteome (Koller et al., 2002), and metabolome (Fiehn et al., 2000) are well under way; however, the study of the ionome, in contrast, is still in its infancy (Lahner et al., 2003; for review, see Hirschi, 2003; Rea, 2003), with the majority of genes and gene networks involved in its regulation still unknown. Moreover, because the ionome is involved in such a broad range of important biological phenomena, including electrophysiology, signaling, enzymology, osmoregulation, and transport, its study promises to yield new and significant biological insight. An understanding of the ionome and how it interacts with other cellular systems such as the genome, proteome, metabolome, and environment are integral to our full understanding of how plants integrate their organic and inorganic metabolisms.

Lahner and colleagues first described the ionome to include all the metals, metalloids, and nonmetals present in an organism (Lahner et al., 2003), extending the term metallome (Outten and O’Halloran, 2001; Williams, 2001; Szpunar, 2004) to include biologically significant nonmetals such as nitrogen, phosphorus, sulfur, selenium, chlorine, and iodine. It is important to note here that the boundaries between the ionome, metabolome, and proteome are blurred. Compounds containing the nonmetals phosphorus, sulfur, or nitrogen, for example, would fall within both the ionome and metabolome, and metals such as zinc, copper, manganese, and iron in metalloproteins would fall within the proteome, or metalloproteome as it has been described (Szpunar, 2004). The elements to be measured in the ionome will be determined by their biological importance or environmental relevance, in conjunction with their amenability to quantitation. However, each element measured must be present in sufficient concentrations in the plant tissue so as to be well above the Limit of Quantitation, defined as the concentration equal to the signal from a blank plus 10 sds of the blank signal.

CHARACTERIZATION OF THE PLANT IONOME: A SINGLE ION AT A TIME

Remarkable progress has been made in describing and understanding the basic biology of nutrient ion homeostasis in plants since its establishment as a scientific disciple in the 19th century (Marschner, 1995). The development and application of modern molecular genetic techniques and completion of the Arabidopsis and rice genomes has accelerated progress; however, much remains to be discovered. During evolution, the first proto cells faced a major obstacle. The outer membrane, while needed to keep the cellular contents organized as a functional unit, created a barrier that prevented the uptake of nutrient ions. As multicellular and terrestrial organisms evolved, the challenges of moving solute ions from the environment to the appropriate tissues further increased. Thus, one of the key advances enabling organisms to survive was the evolution of ion transport systems. Because of their central importance, ion transporters have been the primary focus of most work involved in characterizing the ionome in plants. In the past few years, transporters for many different ions have been characterized (Mäser et al., 2001). Multiple genes, and even multiple gene families, appear to be responsible for transport. This is not surprising considering that transport across different membranes is required in plant tissues with diverse nutritional and energy requirements. In addition, multiple membrane proteins may be needed for ion uptake from the soil to adapt to varying extracellular conditions and nutrient availability. Such paradigms are exemplified in our current understanding of the regulation of numerous mineral ions in plants, including iron, zinc, sodium, phosphorus, potassium, and calcium (Rausch and Bucher, 2002; Sanders et al., 2002; Curie and Briat, 2003; Very and Sentenac, 2003; Zhu, 2003). Though extensive progress has clearly been made, a careful analysis of the Arabidopsis genome reveals the existence of approximately 1,000 ion transporters, most of which have not yet been characterized. Furthermore, it

---

1 This work was supported in part by the National Science Foundation (NSF) Plant Functional Genomics program (grant no. 0077378–DBI).

* E-mail dsalt@purdue.edu; fax 765–494–0391.

is estimated that 5% of the approximately 25,000 genes in the Arabidopsis genome are involved in regulating the ionome (Lahner et al., 2003). Clearly, a major challenge to understanding the genes and gene networks involved in ion homeostasis in plants is to design ways to probe gene function on a genomic scale.

CHARACTERIZATION OF THE PLANT IONOME: MULTIPLE IONS AT A TIME

Progress in our understanding of plant mineral nutrition has been closely related to the development of chemical quantification methods and availability of pure reagents. This relationship is clearly reflected in the timeline of discovery for the essentiality of various micronutrients in plants from iron in 1860 to nickel in 1987. More recent developments in parallel and massively parallel analytical techniques have played a critical role in the explosive development of genomic biology. The advent of DNA microarray technology has certainly accelerated the pace at which genes regulated by ionic changes can be identified. Not surprisingly, many genes are transcriptionally responsive to changes in nutrient availability, including transporters, transcription factors, and signaling factors (Thimm et al., 2001; Negishi et al., 2002; Maathuis et al., 2003; Wang et al., 2003; Wintz et al., 2003). It is clear from these and other studies that plants can respond specifically to availability of individual nutrients and nutrient deficiencies, suggesting that many regulatory pathways exist. The challenge is how to integrate alterations in transcription with a functional understanding of how ion homeostasis networks operate. As part of this challenge, we have developed a high throughput ion-profiling strategy based on the use of recently developed robust inductively coupled plasma mass spectroscopy (ICP-MS) technology for genomic scale profiling of nutrient and trace elements (Lahner et al., 2003; Fig. 1). Using this system, we have analyzed shoot tissue from more than 40,000 plant samples to date, averaging 1,000 per month, including various ecotypes and both fast neutron and T-DNA insertional mutants. We have also analyzed the seed ionome in numerous mutants and ecotypes. During this project, we have also collected data on the shoot and seed ionome of Arabidopsis, under our standard soil growth conditions (Fig. 1).

HIGH-THROUGHPUT ION PROFILING

For comparative ionomics on a genome-wide scale, wild-type and mutant plants are ideally grown side-by-side under identical conditions, in a uniform growth media. They all grow to the same size and are harvested at the same time, sampling equivalent amounts and parts of tissue in every case. They are grown and harvested under clean room conditions, using tools that won’t impart any measured elements to the samples, washed of any surface contaminants, dried, weighed, and analyzed accurately and precisely. The data are processed efficiently, summarized in an easily understood format, and made widely accessible. However, in the imperfect world, a screen of a significant portion of the genome takes months or years. Conditions change, personnel change. Samples get contaminated with growth media, growth media varies from batch to batch. Plants grown on soil are under- or overwatered. Sample sizes vary and include different tissues from different plants. Instruments vary with the maintenance cycle and operating conditions. Data are mixed up, lost, or misinterpreted, and programs have bugs. The success or failure of such an ionomic project is determined by which of these two scenarios we stand closer to. Within these boundaries lies another dimension to consider: sample throughput versus the quality and breadth of the data collected. The best approach to this key tradeoff is by no means universally agreed upon. Nearer to one end of this dimension lie scientists such as Sydney Brenner, who feel that “...data that goes into a database should be complete, accurate and permanent, so you never have to do it again” (Duncan, 2004, p. 21). At the other end lie scientists who emphasize speed to maximize the number of mutants found, hoping to find the lower-hanging fruit by screening larger portions of genomes, and tend to rely on various statistical tools to extract good information from noisy and incomplete data.

For ionomic analysis samples are typically digested in concentrated acid and diluted before analyzing. Nitric acid digests most plant material easily and, of the common inorganic acids, interferes with ICP-MS analysis the least. Open-air digestion below the boiling point works well, while microwave digestion is becoming more common, especially where loss of an analyte of interest is a concern. Digestion time is much shorter under the higher pressure and temperatures of the microwave digestion apparatus; however, capacity issues shift the overall speed equation toward open-air digestion, where several hundred samples can be run in the same hood that would vent the microwave, and with far less sample handling. The three most common methods for elemental analysis are atomic absorption spectroscopy, ICP-optical emission spectroscopy (ICP-OES), and ICP-MS. Atomic absorption spectroscopy quantitates one element at a time or a few elements in rapid succession. This method is very well established and quite precise, but not nearly as sensitive as ICP-MS and with a dynamic range of only 3 or 4 orders of magnitude. ICP-OES, often referred to as simply ICP, is less sensitive than ICP-MS; however, some of this sensitivity is won back by the robustness of ICP in more concentrated matrices. ICP is a reasonable choice for an ionomics screen, at the possible expense of some of the trace elements due to its lower sensitivity. Both ICP and ICP-MS can measure multiple elements essentially simultaneously in the same sample, a very
important property in an ionomics project. One critical advantage of ICP-MS is that it allows for a smaller sample size due to its greater sensitivity, making nondestructive sampling of small plants possible, a prerequisite in a random forward genetic screen—interesting mutants need to be saved not destroyed by the ICP. ICP-MS also has a dynamic range of 6 or 7 orders of magnitude, making it possible to simultaneously measure both macronutrients and micronutrient concentrations in the same sample. Effective genome-scale ionomics requires that many thousands of plant samples will need to be grown, harvested, dried, weighed, digested, and analyzed over long periods of time. Such an effort means that hundreds of samples will need to be processed in a consistent manner, weekly, for several years. Such an effort continuously generates thousands of pieces of data, and the final critical task in ionomics is data handling, including everything from getting the data off the analytical instrument, through data tracking, data storage, data processing, analyses, and publication.

Analysis of raw data generated in an ionomics project can be carried out in a number of valid ways. The key objective is identification of plants with disturbed ionomes. This goal does not require that actual concentrations of any of the elements be determined. Instead, interesting mutants can be identified by their overall elemental profiles using average signal normalization and various supervised and unsupervised clustering systems such as discriminant analysis, principle component analysis, and neural net trolling. Clearly, an efficient and robust process containing all these components needs to be designed and implemented before any successful ionomic project can be performed (Lahner et al., 2003).

**LINKING THE IONOME AND GENOME**

By considering the ionome as a whole, the concept of ion homeostasis networks arises, in which various ions within an organism are coordinately regulated.
The observation that only 11% of the 50 Arabidopsis ion-profile mutants recently identified (Lahner et al., 2003) showed changes in only one element strongly supports the existence of such regulatory networks in the ionome of plants. Characterization and mapping of these ion homeostasis networks should help uncover not only their genetic basis but also how they interact with both the proteome and the metabolome. Identification of genes underlying any biological phenom-
ena can take the two different but complementary approaches of forward and reverse genetics. Forward genetics is the more traditional approach of mapping genotypic variation to a specific phenotype. Genotypic variation can be either naturally occurring, such as between different ecotypes of Arabidopsis (Alonso-Blanco and Koornneef, 2000), or induced using various mutagens including ethyl methanesulfonate (EMS), x-rays, fast neutrons (FN; Koornneef et al., 1982), T-DNA (Azpiroz-Leehan and Feldmann, 1997; Krysan et al., 1999), and transposable elements such as Disso-
ciation (Parinov et al., 1999). Once a plant population has been established with significant genotypic variation, a suitable screen needs to be developed to identify plants with the phenotype of interest. The probability of identifying a plant harboring a mutation in a gene that affects the trait of interest, in this case the ionome, is dependent on both the mutation frequency and size of the gene(s). Mutation frequency varies between muta-
gens, with FN and EMS producing on average 30 to 60 mutations per diploid genome (Koornneef et al., 1982), compared to 1.4 mutations for T-DNA (Feldmann, 1991). To perform a saturation screen using an EMS or FN mutagenized population would therefore require phenotyping of approximately 10,000 to 20,000 M₂ plants, whereas the same screen with T-DNA would require 200,000 to 400,000 M₃ plants. Clearly, even when using an EMS or FN mutagenized population, the screening system used to identify plants with an altered ionome needs to be relatively high throughput in order to achieve saturation.

Achieving high-throughput ICP-MS analysis at the high precision needed to produce a viable screen of the ionome is challenging and requires both good analyt-
tical techniques and data handling. We have performed such a screen on shoot tissue from an Arabidopsis FN mutagenized population of approximately 6,000 M₂ plants and identified 51 mutants with altered shoot ionomes (Lahner et al., 2003). Of these mutants, only one mutant, frd3-5, appeared to show dominance after analysis of the high manganese phenotype in the M₃ generation. However, to confirm the recessive nature of the majority of these mutations backcrossing to the wild type (Col 0), selfing of the hybrid F₂ and scoring of the ionomic phenotype in the F₃ is required. Such analyses have been performed on three of these ionomics mutants (145:01, 121:33, and 132:31) that show alterations in various elements, including boron, calcium, potassium, molybdenum, sodium, and nickel, and all were found to be recessive. A similar ICP-MS based forward genetic screen is also under way in the laboratory of Mark Tester (http://www. designscene.com.au/clients/n3/30_group/tester_mark.htm). This project plans to screen activation tagged lines of Arabidopsis to identify genes involved in regulating the ionome; however, few details are yet available. Activation tag mutagenesis involves transforma-
tion with T-DNA vectors that contain a multi-
erized transcription enhancer (Weigel et al., 2000). Such a strategy allows for the identification of both insertional mutants and dominant gain-of-function mutations, in which the transcription enhancer activates expression of genes adjacent to the insertion of the T-DNA. This can be contrasted with other forms of mutagenesis, including EMS and FN, in which loss-of-
function mutations are more common.

The availability of Arabidopsis high-density gene arrays now makes it possible to simultaneously genotype plants for several hundred thousand loci. By using total genomic DNA instead of mRNA for hybrid-
ization and pooling DNA from only 15 homozygous recombinants displaying the mutant phenotype, it is possible to map a locus to approximately 12 cM (Borevitz et al., 2003). Simulations suggest that using DNA from a pool of >200 plants would allow map-
ing down to <0.5 cM (Hazen and Kay, 2003). Because FN mutagenesis produces deletions of between 2 to 4 kb (Li et al., 2001) and assuming the deletion is within an exon, bulk segregant analysis of pooled F₂ plants using genomic DNA as a probe should also provide a rapid strategy for the identification of mutant genes. The emerging high-density Arabidopsis genome tiling arrays should greatly facilitate the identification of genomic polymorphisms, allowing the rapid identifi-
cation of deletions produced by FN mutagenesis in a single F₂ bulk segregant analysis experiment. Such an approach could revolutionize forward genetics in Arabidopsis, making it possible to identify a mutation within 3 to 4 months. Positional cloning of the genes responsible for ion-profile changes in several of the ionomic mutants (Lahner et al., 2003) is under way in our laboratory. Outcrossoes to Landsberg erecta (Ler) have been performed, hybrid F₁ allowed to self, and the F₂ progeny screened for homozygous mutants. These F₂ mapping populations are now being screened for cosegregating genetic polymorphisms.

As an alternative to laboratory-induced mutations, genetic variation occurring among and within natural populations of Arabidopsis can be used (Alonso-Blanco and Koornneef, 2000). Since Arabidopsis shows a wide geographic distribution, many Arabidopsis ecotypes or accessions have been collected and are available from the Arabidopsis Biological Resource Center (ABRC). Considerable variation for such traits as resistance to biotic and abiotic stress, development, and metabolic traits has been described (for review, see Alonso-Blanco and Koornneef, 2000). Observed variation between accessions can either be qualitative, defined by phenotypic distributions that fall into discrete classes and is caused by one or two major loci, or quantitative, defined by a continuous pheno-
typic distribution and caused by the combined effect of multiple loci. Natural variation in Arabidopsis seed and shoot phosphate accumulation is known to exist between the Ler and Cvi accesses (Bentsink et al., 2003), and for potassium, sodium, calcium, magnesium, iron, manganese, zinc, and phosphorus in seeds of numerous ecotypes (Vreugdenhil et al., 2004). Analyses of Ler/Cvi recombinant inbred lines revealed quantitative trait loci (QTL) that explain between 10% and 79% of this variation for the different elements (Vreugdenhil et al., 2004). Natural variation in several Arabidopsis ecotypes has also been observed for shoot caesium. Four QTL accounting for >80% of the genetic contribution governing this variability have been identified in Ler × Col recombinant inbred lines, in which this variability has been fixed (Payne et al., 2004). The robust and precise quantification of the ionomic variation in various ecotypes holds the promise of Mendelian or QTL mapping to identify genes involved in regulating the ionome in Arabidopsis.

An alternative to the forward genetic approach described is the opposite strategy of starting with a mutation in a known gene and asking the question, “Does this mutation have an ionomic phenotype?” Such an approach switches the focus from one of high-throughput screening and mapping to an approach that requires in-depth biochemical and physiological analyses of the mutant. A close-to-saturation collection of Arabidopsis T-DNA insertion mutants with sequenced boarders is curated at the Salk Institute Genomic Analysis Laboratory (SIGnAL), making this approach attractive in Arabidopsis. High-throughput screening for induced point mutations (TILLING) also makes it possible to identify alternative alleles in genes of interest in numerous species, including maize (Zea mays), black cottonwood (Populus balsamifera subsp. trichocarpa), Brassica oleracea, and Lotus japonicus. The use of genomic DNA pooling and PCR also makes the identification of mutants in specific genes possible in other species using fast neutron deletion mutagenesis (Li et al., 2001). As part of the ionomics project in our laboratory, we are in the process of quantifying the shoot ionome in approximately 1,000 homozygous T-DNA insertional lines of Arabidopsis. Mutants have been chosen with insertions in exons of genes thought to be involved in regulating the ionome, including ion transporters and various regulatory proteins. To date, we have analyzed the shoot ionome in 12,000 plants, representing insertions in approximately 600 genes.

**FUTURE DIRECTIONS**

To date, ionomics has been applied to bulk tissue samples. Such analyses only provide a very limited view of the tissue, and cellular and subcellular complexities of ion homeostasis mechanisms. The ability to profile the elemental content of different plant tissues such as meristematic and vascular tissue requires a 10- to 50-μm spatial sampling resolution. Such imaging has been achieved in vivo for individual elements such as selenium in plants using x-ray spectroscopy (Pickering et al., 2000, 2003), though not for multi-element analyses using ICP-MS. However, use of laser ablation sampling coupled to ICP-MS holds promise for development of high resolution ionomic imaging (Narewski et al., 2000; Kang et al., 2004). If developed in living plant tissue, such technology would open up a completely new window onto the ionome, allowing changes in the total shoot, seed, or root ionomes, for example, to be mapped to specific tissues and cell types. Such ionomic imaging would also allow colocalization of in vivo gene expression and protein localization patterns with ionomic changes, providing spatial linkage between gene, protein, and ionomic function.

**COMMUNITY IONOMIC RESOURCES**

To maximize the value of any large-scale genomics effort, it is critical that the data be made available to as wide a selection of people as possible. To facilitate such a process, we have developed a searchable online database containing ionomic information on more than 22,000 plants. The database can be searched for fast neutron mutagenized mutants altered in a specific or set of elements, as well as on gene name and AGI gene codes for reverse genetics in T-DNA insertional lines. This ionomics database can be found at http://hort.agriculture.purdue.edu/Ionomics/database.asp. Ionomic mutants reported by Lahner et al. (2003) are detailed in the ionomics database and available through the ABRC (http://www.biosci.ohio-state.edu/approximately plantbio/Facilities/abrc/abchio.htm). Seeds of other M₄ fast neutron mutant lines identified in the database are available on request (dsalt@purdue.edu), dependent on availability. Genes that appear to alter the ionome, identified in the T-DNA insertional ionomics database, can be further characterized using T-DNA insertional lines curated at the SIGnAL database (http://signal.salk.edu/cgi-bin/tdnaexpress). The Purdue Ionomics Database will be periodically updated and expanded, and our ultimate goal is to provide a biologist-friendly synthetic laboratory that integrates ionomic data with other types of genomic, biochemical, and physiological datasets. Such an environment will allow in silico experiments to be performed that will facilitate the development of sophisticated hypotheses that can be tested with wet experiments. It is hoped that such an environment will accelerate the speed at which researchers design and perform experiments that deliver new and significant biological insight.

**ACKNOWLEDGMENTS**

The development of ionomics as a functional genomics approach has been supported in part by the NSF Plant Functional Genomics program (0077378–DBI) through a collaborative grant to Mary Lou Guerinot, David Eide, Jeff
Harper, David E., Salt, Julian Schroeder, and John Ward. We also thank Brett Labher for his insightful and sustained input on the design, implementation, and running of our omics system at Purdue and help with the content of this update.

Received June 7, 2004; returned for revision July 9, 2004; accepted July 12, 2004.

LITERATURE CITED


