

### The omicization of agrochemical research

In 1920 Hans Winkler, Professor of Botany at the University of Hamburg, Germany combined the words gene and chromosome to make the word “genome.”<sup>1)</sup> The suffix “ome” originated from the word chromosome, which comes from the Greek stems “color” and “body.” Informally, the neologism *omics* has come to refer to a comprehensive study involving the acquisition of vast data sets. An omics approach can be considered to be large-scale data-rich biology consisting of a heavy data-mining or bioinformatics component. The modern concept of omics was initiated by the Human Genome Project, which was launched in 1986 and essentially completed in April 2003. The sequence of the last remaining chromosome was published in May of this year,<sup>2)</sup> giving a complete genome at a cost of roughly US\$3 billion. This trend has been continued with the Human Metabolome Project, a US\$7.5 million project launched in January 2005 and funded by Genome Canada.<sup>3)</sup> The project mandate is to identify, quantify, catalogue and store all metabolites that can potentially be found in human tissues and biofluids at concentrations greater than one micromolar. The Human Proteome Organization (HUPO) is also hard at work on the Human Proteome Project to identify the components of the human plasma, liver and brain proteomes.<sup>4)</sup> Similar projects are underway in many other species, including mammals, plants, and environmentally relevant indicator species as well as bacteria.

The agrochemical industry is one of the world's most research-intensive sectors, with the top 10 agrochemical companies spending US\$2.5 billion, or 7.5% of sales, on research and development (R&D) in 2004.<sup>5)</sup> The most significant area of the R&D process is new product discovery and development, with costs representing 31.3% of overall R&D expenditures.<sup>5)</sup> In 2004, the two most significant components of agrochemical product discovery were chemical synthesis and laboratory biology, accounting for 51.1% of the overall cost.<sup>5)</sup> These expenditures were exasperated by significant decreases in the success rate in the production of novel agrochemical commercial products. In the year 2000, it was necessary to test roughly 140,000 substances to obtain a commercial product as opposed to only 20,000 compounds in 1980.<sup>6)</sup> One approach to address these shifting research trends is through the application of omics technologies.<sup>7)</sup> In agrochemistry, the use of genomics to identify suitable product candidates and the introduction of combinatorial chemistry to produce compounds for biological screening were identified as important developments over the last 10 years.<sup>5)</sup> For example, it is expected that functional genomics will be useful in the identification of novel targets for herbicide development and aid in the development of assays systems for discovering lead compounds.<sup>8)</sup> Towards this end, commercial gene chips are available for a number of important commodity crops including: barley, citrus, grapes, maize, rice, soybean, sugar cane, tomato and wheat.<sup>9)</sup> The ability to monitor wide-scale fluctuations in

gene expression following experimental treatment can be extremely useful for identifying new lead compounds as well as modes of action. For example, the GeneChip Arabidopsis ATH1 Genome Array contains more than 22,500 probe sets representing approximately 24,000 genes.<sup>9)</sup> Lechelt-Kunze *et al.* used an earlier version containing only 8247 genes to identify the mode of action of the herbicides flufenacet and benfuresate.<sup>10)</sup> Their results pointed to the inhibition of very-long-chain fatty acid elongase (VLCFAE) as the source of the observed mutant fiddlehead phenotype. In addition to genomics methods, proteomic and metabolomic technology platforms are advancing rapidly making routine use practical and affordable. Bioinformatic resources are expanding equivalently with the Kyoto Encyclopedia of Genes and Genomes (KEGG) data base currently listing sequence information and subsequent metabolic pathway maps for 35 eukaryotes, 322 bacteria and 27 archaea.<sup>11)</sup>

Gene expression profiling (genomics) and metabolic profiling (metabolomics) are becoming an increased utility in agrochemical research settings.<sup>7)</sup> Used as fingerprint methods, these enable a fast and reliable detection of known modes of action and a clear identification and classification of unknown modes of action. For example, artificial neural network analysis of <sup>1</sup>H NMR spectra was used to classify changes in total metabolic profile in aqueous plant extracts of corn (*Zea mays*) caused by herbicide application.<sup>12)</sup> This method was also used by Ott *et al.* to detect and classify 19 different distinct modes-of-action in corn following herbicide treatment.<sup>13)</sup> Another important application of omics technologies in the agricultural sciences is the analysis of the potential effects of genetic modifications upon metabolite composition. There is consumer concern over the effects of transgene insertion upon the “nutritional content” of food crops. These concerns could be partially addressed through comprehensive metabolic profiling of the plant pre- and post-modification. For example, metabolic profiling was used to show that genetically modified potatoes did not exhibit changes in metabolite composition.<sup>14)</sup> The metabolic profiles of five transgenic cucumber lines were compared taking into consideration their transgene integration sites, with each location shown to possess a specific metabolic profile.<sup>15)</sup> A proteomic approach identified changes in gluten proteins in genetically modified durum wheat<sup>16)</sup> as opposed to studies in the virus-resistant tomato,<sup>17)</sup> which evidenced no alterations in the expression of major proteins following genetic manipulation. The prevalence of transgenic crops is expected to increase as their appeal is expanded with a wider variety of modifications and crop varieties. For example, transgenic rice has been developed that is capable of growing in the presence of the majority of currently available commercial herbicides,<sup>18,19)</sup> which could be extremely useful throughout large parts of the world. As the introduction of transgenic commodity crops continues to increase, it will be important to have established methods for verifying any potential differences in protein and metabolite production. These points could be par-

ticularly important from the standpoint of monitoring for allergen development and nutritional quality and assist greatly in addressing public concerns regarding safety issues.

The sequencing of a number of important insect genomes including the flies *Drosophila melanogaster*<sup>20)</sup> and *Drosophila pseudoobscura*,<sup>21)</sup> the mosquito *Anopheles gambiae*,<sup>22)</sup> the honey bee *Apis mellifera*,<sup>23)</sup> the pea aphid *Acyrtosiphon pisum*,<sup>23)</sup> three species of parasitic wasp: *Nasonia vitripennis*, *Nasonia giraulti*, and *Nasonia longicornis*,<sup>23)</sup> and the silk worm *Bombyx mori*<sup>24)</sup> as well as the accumulation of genomic and mRNA expression data for a number of additional insects provides a fertile background for comparative genomics. In addition, post-genomic research aiming to survey gene and protein expression patterns in various developmental and physiological contexts will complement these studies and greatly expand our knowledge of insect biology. For example, comparative genomics has been used to explore the role of insect transferrin and ferritin proteins, which are important proteins in iron metabolism that potentially play a role in insect defence against oxidative stress and infection.<sup>25)</sup> The sequencing of the *Drosophila melanogaster* genome has expanded our knowledge of how insecticides act on molecular targets in the insect nervous system, providing evidence that specific receptors and ion channels are targeted by distinct chemical classes of insect control agents.<sup>26)</sup> Pedra *et al.* employed a proteomics approach to examine differences in DDT metabolism-based resistance in *Drosophila melanogaster*.<sup>27)</sup> Proteomic profiling was performed in two DDT susceptible phenotypes and three DDT resistant lines, with results suggesting that insecticide resistance impacts different metabolic pathways in *Drosophila* phenotypes selected with the same pesticide. Selkirk *et al.* examined the molecular diversity and tissue distribution of acetylcholinesterases in the nematode *Caenorhabditis elegans* and proposed that functional genomics could be used to develop more selective inhibitors.<sup>28)</sup> The southern cattle tick, *Boophilus microplus* (Canestrini) has been proposed for sequencing as a method to develop novel control technologies and overcome current problems of pesticide resistance.<sup>29)</sup> For example, a genome could provide information on genes involved in the regeneration of amputated tick limbs and transitions through developmental stages, mechanisms which currently are largely unknown. Taken together, these broad-based omics approaches could potentially address one of the current main issues in agrochemistry, a lack of novel targets.<sup>5,7)</sup>

One of the initial areas of fruitful research has been in the application of proteomics to examine the effects of herbicide treatment upon target organism proteins. For example, treatment of the grapevine (*Vitis vinifera* L.) with the herbicide flumioxazin resulted in 33 distinct proteins displaying a differential expression pattern.<sup>30)</sup> These proteins included a diverse range of functions including photosynthesis-related proteins and antioxidant systems, providing information on systemic effects of herbicide application. Application of the her-

bicide safener fluxofenim upon the wheat *Triticum tauschii* caused 18 proteins to be induced, including 15 glutathione S-transferase (GST) subunits as well as 3 proteins showing homology to the aldo/ketose reductase family and with proteins having known roles in glycolysis and the Krebs cycle.<sup>31)</sup> Herbicide safeners were also showed to induce GSTs as well as glucosyltransferases in maize and Arabidopsis.<sup>32)</sup> Teixeira and coworkers used a proteomics approach to differentiate the effects of 2,4-dichlorophenoxyacetic acid (2,4-D) treatment upon *Saccharomyces cerevisiae*.<sup>33)</sup> These selected papers represent a small selection of the types of studies that can be performed with omics applications in agricultural research. It is expected that the breadth of the applications as well as the analytical platforms will continue to expand at a rapid rate, providing for increased throughput and ease of use.

This issue of the Journal of Pesticide Science highlights a number of applications of omics-related technologies and methods in order to illustrate an important trend in agrochemical research. To stress the nature of the expanding research field, non-traditional applications are emphasized to illustrate other potential research directions for the study of agrochemical development, metabolism and environmental fate. Tietjen *et al.* have reviewed the applications of omics technologies to the conventional agrochemical development method.<sup>7)</sup> The influence of omics methods can already be observed at the IUPAC meeting in Kobe this summer with the theme “Evolution for Crop Protection, Public Health, and Environmental Safety.” Platform sessions include: Genomics, Proteomics and Metabolomics” and “Drug Design Based Upon Agro-genomics” with a presentation on a “Proposal for SAR-omics as a paradigm for lead evolution in drug design.”<sup>34)</sup> As the field of agrochemical research changes with advances in technology, it will be important to take advantage of the potential that omics-applications offer. This issue begins with a review by Lin *et al.* who discuss the use of metabolomics to examine the effects of organism exposure to xenobiotic or environmental stressors. An important component of agrochemical research involves evaluating the potential effects of non-target organism exposure to pesticides. For example, this approach has been used to identify significant metabolic perturbations to the early life stages of Chinook salmon exposed to organophosphate and pyrethroid insecticides.<sup>35)</sup> Due to the ability of NMR-based metabolomics to detect a wide range of biological molecules, it is expected to be a useful approach to evaluate effects of xenobiotic exposure.<sup>36)</sup> In addition, whole organism nondestructive studies can be performed.<sup>37)</sup> This approach can be complemented with genomics studies as described by Garcia-Reyero and Denslow who discuss applications of ecotoxicogenomics to studies in fish following exposure to organochlorine insecticides. Ecotoxicogenomics involves the study of gene and protein expression in non-target organisms in response to environmental toxicant exposures.<sup>38)</sup> It is expected that ecotoxicogenomic tools will assist in overcoming some of the key challenges in ecotoxicology includ-

ing predicting toxicant responses across the diversity of phylogenetic groups present in an ecosystem, estimating how changes in one ecological or trophic level affect other levels and predicting the influence of time-varying exposures upon toxicant responses.<sup>38,39)</sup> A significant advantage of these approaches in studying xenobiotic effect(s) upon organism health is the breadth of the analysis. A genomic study combined with a metabolomics approach will provide a comprehensive picture of the effect of exposure upon an organism's biology. This thoroughness enables multiple conclusions to be drawn from a single well-designed study, potentially reducing expenditures in the long run. Studies focused on fish systems are particularly important as they are extensively used to investigate ecosystem health and xenobiotic toxicity. Sequencing efforts have been completed or are currently underway for several fish species, including zebrafish (*Danio rerio*),<sup>40)</sup> medaka (*Oryzias latipes*),<sup>41)</sup> pufferfish (*Fugu rubripes*)<sup>42)</sup> and *Tetraodon nigroviridis*)<sup>43)</sup> and the stickleback (*Gasterosteus aculeatus*).<sup>44)</sup> A number of different omics efforts are being placed upon the fathead minnow (*Pimephales promelas*), which has become a standard model of aquatic toxicology.<sup>45)</sup> The U.S. Department of Energy in conjunction with the Environmental Protection Agency has developed a fathead minnow cDNA library that includes more than 250,000 expressed sequenced tags. In addition, both proteomic and metabolomics efforts are currently underway for the fathead minnow.<sup>45,46)</sup> Data from these endeavors can potentially lead to a full understanding of fathead minnow systems biology, enabling for rapid estimations of xenobiotic effects upon the entire organism. As the fathead minnow is a standard U.S. EPA test species, these tools could be extremely valuable. One could envision an EPA test package that included microarrays for the 3 standard test species (the phytoplankton *Selenastrum capricornutum*, the zooplankton *Ceriodaphnia dubia* and the fathead minnow) to provide a quick and efficient estimation of effects upon ecosystems following xenobiotic exposure or other stressors. This method could be especially useful in diagnosing the effects of multiple stressors.

Another extremely important research area in agrochemical research is insect vector control to reduce the spread of infectious diseases, with the IUPAC meeting devoting a symposium to "Control Agents for Vectors of Communicable Diseases."<sup>34)</sup> The past few years has seen a tremendous rise in the amount of interest, and more importantly funds, dedicated to the control of tropical diseases such as malaria. This research has been targeted to the discovery of novel pesticides for conventional vector control as well as the generation of transgenic organisms incapable of hosting/spreading the parasite.<sup>47,48)</sup> Given the scale of the impact of malaria on human suffering, with 1–2 million deaths annually, research focusing on eradication of the disease is of extremely high priority.<sup>49)</sup> The agrochemical research community is in a position to make a significant impact in this research area by providing novel ways to manage the vector. A wealth of genomics infor-

mation has been provided with the sequencing of all 3 components of the malaria lifecycle (parasite, vector and host) as well for a number of different apicomplexan parasites.<sup>50,51)</sup> These data should be useful for comparative genomics studies that aim to identify vital metabolic pathways that are selective for the target. In this issue, Hayes and coworkers critically analyze the benefits of omics-related technologies to the malaria research field on a method-by-method basis. Their work, while focusing on malaria, is equally applicable to other infectious diseases and provides a useful overview of applications of omics technologies to the study of disease control.

One of the greatest shifts in the research paradigm for the bench scientist is the need for bioinformatics knowledge. Omics technologies are by definition data intensive and require advanced analysis methods dependent upon bioinformatic approaches. Subsequently, a focus of this thematic issue is a number of newly developed software packages designed to analyze omics-related data sets. These packages are all similar in their approach in that they are designed as tools to be used by the bench researcher to assist in the analysis of large-scale data sets across multiple analytical platforms. A key component is the drive towards an integrative systems biology approach, incorporating omics data from multiple studies. Kadowaki *et al.* provide a bridge between more traditional agrochemical-based QSAR research approaches to bioinformatics with their review on a novel classification method. They use endocrine disruptors as an example to present novel methods to incorporate genomics data with metabolic pathway maps in order to study degradation processes. This work is followed by a paper by Arakawa and Tomita who discuss their G-language Genome Analysis Environment, which is a generic genome analysis program that constructs an integrated environment for the development of analysis software. The remainder of the review issue consists of 3 short commentaries designed to succinctly present a few of the plethora of available bioinformatics tools. Klukas *et al.* present a software tool called VANTED (Visualisation and Analysis of NeTworks containing Experimental Data), which aims to provide visual exploration and statistical analysis of complex biochemical data sets through the integration of genome, proteome and metabolome data. Sakurai and Shibata present the web-based KaPPA-View (Kazusa Plant Pathway Viewer) that was developed to represent quantitative data for individual transcripts and/or metabolites on plant metabolic pathway maps. Aoki-Kinoshita provides an overview of the KEGG resource, which is designed to enable the computational prediction of higher-level complexity of cellular processes and organism behaviour from genomic and molecular information. These expert reviews and commentaries provide an overview of a small fraction of the bioinformatics tools available to researchers. All of these resources can be accessed free of charge by the academic research community. Interested readers are highly encouraged to explore these and other methods and to provide feedback to the developers in order to continu-

ally improve the tools available to the research community. It is vital that the bioinformatics and bench research communities work together to iteratively develop applications that forward the research field. This issue ends with a commentary by Bamba and Fukusaki on plant metabolomics that discusses a number of key issues involved in experimental design.

Biology has increasingly become a data-rich subject involving large-scale data sets that are mined for novel insight. The continued evolution of agrochemistry depends upon the adoption of novel methods to target discovery, mode of action and lead compound identification. The use of omics technologies is a logical approach to expanding the arsenal of tools available in this important industry. As we increase our repertoire of tools with which to explore and expand research into agrochemicals and biotechnological approaches to the agricultural sciences, it is vital that we embrace emerging technologies and explore new research possibilities. It is possible that along the way, we will create our own approach of “agronomics.”

Craig E. Wheelock and Hisashi Miyagawa  
Journal of Pesticide Science

### References

- H. Winkler: “Verbreitung und Ursache der Parthenogenesis im Pflanzen- und Tierreiche,” Verlag Fischer, Jena, 1920.
- S. G. Gregory, *et al.*: *Nature* **441**, 315–321 (2006).
- <http://www.metabolomics.ca/>
- <http://211.32.65.137/hpp/hppp.htm>
- P. McDougall: “Agrochemical Industry Research and Development Expenditure,” Phillips McDougall, Midlothian, UK, p. 23, 2005.
- P. McDougall: “The Cost of New Agrochemical Product Discovery, Development and Registration in 1995 and 2000,” Phillips McDougall, Midlothian, UK, p. 24, 2003.
- K. Tietjen, M. Drewes and K. Stenzel: *Comb. Chem. High T. Scr.* **8**, 589–594 (2005).
- W. Lein, F. Bornke, A. Reindl, T. Ehrhardt, M. Stitt and U. Sonnewald: *Curr. Opin. Plant Biol.* **7**, 219–225 (2004).
- <http://www.affymetrix.com/products/arrays/index.affx>
- C. Lechelt-Kunze, R. Meissner, M. Drewes and K. Tietjen: *Pest Manag. Sci.* **59**, 847–856 (2003).
- M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki and M. Hirakawa: *Nucleic Acids Res.* **34**, D354–357 (2006).
- N. Aranibar, B. Singh, G. Stockton and K. Ott: *Biochem. Biophys. Res. Commun.* **286**, 150–155 (2001).
- K. Ott, N. Aranibar, B. Singh and G. Stockton: *Phytochemistry* **62**, 971–985 (2003).
- G. Catchpole, M. Beckmann, D. Enot, M. Mondhe, B. Zywicki, J. Taylor, N. Hardy, A. Smith, R. King, D. Kell, O. Fiehn and J. Draper: *Proc. Natl. Acad. Sci. USA* **102**, 14458–14462 (2005).
- N. Tagashira, W. Plader, M. Filipecki, Z. Yin, A. Wisniewska, P. Gaj, M. Szwacka, O. Fiehn, Y. Hoshi, K. Kondo and S. Malepszy: *Cell. Mol. Biol. Lett.* **10**, 697–710 (2005).
- A. Di Luccia, C. Lamacchia, C. Fares, L. Padalino, G. Mamone, B. La Gatta, G. Gambacorta, M. Faccia, N. Di Fonzo and E. La Notte: *Ann. Chim.* **95**, 405–414 (2005).
- D. Corpillo, G. Gardini, A. Vaira, M. Basso, S. Aime, G. Accotto and M. Fasano: *Proteomics* **4**, 193–200 (2004).
- S. Hirose, H. Kawahigashi, K. Ozawa, N. Shiota, H. Inui, H. Ohkawa and Y. Ohkawa: *J. Agric. Food Chem.* **53**, 3461–3467 (2005).
- H. Kawahigashi, S. Hirose, H. Ohkawa and Y. Ohkawa: *J. Agric. Food Chem.* **53**, 8557–8564 (2005).
- M. D. Adams, *et al.*: *Science* **287**, 2185–2195 (2000).
- S. Richards, *et al.*: *Science* **298**, 129–149 (2002).
- <http://www.hgsc.bcm.tmc.edu/projects/>
- K. Mita, M. Kasahara, S. Sasaki, Y. Nagayasu, T. Yamada, H. Kanamori, N. Namiki, M. Kitagawa, H. Yamashita, Y. Yasukochi, K. Kadono-Okuda, K. Yamamoto, M. Ajimura, G. Ravikummar, M. Shimomura, Y. Nagamura, I. T. Shin, H. Abe, T. Shimada, S. Morishita and T. Sasaki: *DNA Res.* **11**, 27–35 (2004).
- B. Dunkov and T. Georgieva: *Insect Biochem. Mol. Biol.* **36**, 300–309 (2006).
- V. Raymond-Delpech, K. Matsuda, B. Sattelle, J. Rauh and S. DB: *Invert. Neurosci.* **5**, 119–133 (2005).
- J. Pedra, R. Festucci-Buselli, W. Sun, W. Muir, M. Scharf and B. Pittendrigh: *Proteomics* **5**, 258–269 (2005).
- M. Selkirk, O. Lazari and J. Matthews: *Parasitology* **131**, Suppl: S3–18 (2005).
- F. Guerrero, V. Nene, J. George, S. Barker and P. Willadsen: *J. Med. Entomol.* **43**, 9–16 (2006).
- A. Castro, C. Carapito, N. Zorn, C. Magne, E. Leize, A. Van Dorsseleer and C. Clement: *J. Exp. Bot.* **56**, 2783–2795 (2005).
- Q. Zhang and D. Riechers: *Proteomics* **4**, 2058–2071 (2004).
- R. Edwards, D. Del Buono, M. Fordham, M. Skipsey, M. Brazier, D. Dixon and I. Cummings: *Z. Naturforsch. [C]* **60**, 307–316 (2005).
- M. Teixeira, P. Santos, A. Fernandes and I. Sa-Correia: *Proteomics* **5**, 1889–1901 (2005).
- <http://www.iupac2006.jtbcom.co.jp/>
- M. Viant, C. Pincetich and R. Tjeerdema: *Aquat. Toxicol.* **77**, 359–371 (2006).
- M. Viant, E. Rosenblum and R. Tjeerdema: *Environ. Sci. Technol.* **37**, 4982–4989 (2003).
- M. R. Viant, C. A. Pincetich, J. H. Walton, R. S. Tjeerdema and D. E. Hinton: *Mar. Environ. Res.* **54**, 553–557 (2002).
- J. Snape, S. Maund, D. Pickford and T. Hutchinson: *Aquat. Toxicol.* **67**, 143–154 (2004).
- A. Miracle and G. Ankley: *Reprod. Toxicol.* **19**, 321–326 (2005).
- [http://www.sanger.ac.uk/Projects/D\\_rerio/index.shtml](http://www.sanger.ac.uk/Projects/D_rerio/index.shtml)
- <http://medaka.utgenome.org/>
- S. Aparicio, *et al.*: *Science* **297**, 1301–1310 (2002).
- O. Jaillon, *et al.*: *Nature* **431**, 946–957 (2004).
- <http://www.genome.gov/12512292>
- G. Ankley and D. Villeneuve: *Aquat. Toxicol.* **78**, 91–102 (2006).
- T. Collette, D. Ekman, J. Kenneke, T. Whitehead, D. Villeneuve, M. Kahl, K. Jensen and G. Ankley: *Society for Environmental Toxicology and Chemistry (SETAC)*, (2005) Baltimore, MD.

- 
- 47) C. Hill, F. Kafatos, S. Stansfield and F. H. Collins: *Nat. Rev. Micro.* **3**, 262–268 (2005).
- 48) <http://www.gcgh.org/subcontent.aspx?SecID=412>
- 49) B. M. Greenwood, K. Bojang, C. J. Whitty and G. A. Targett: *Lancet* **365**, 1487–1498 (2005).
- 50) N. M. El-Sayed, *et al.*: *Science* **309**, 409–415 (2005).
- 51) N. M. El-Sayed, *et al.*: *Science* **309**, 404–409 (2005).