European Journal of Plant Pathology **109:** 101–108, 2003.

Mini review

Agricultural genomics: An approach to plant protection

Hinanit Koltai and Hanne Volpin
Department of Genomics, ARO, Volcani Center, Bet Dagan 50250, Israel (Phone: +972 3-9683904/3671; Fax: +972 3-9604180; E-mail: hkoltai@volcani.agri.gov.il)

Accepted 7 November 2002

Key words: bioinformatics, crop, functional genomics, pathogen, plant pathology

Abstract

Genomics studies, focused on whole-genome analysis, have opened up a new era for biology in general, and for agriculture in particular. Along with the use of genetic plant models and the progress in sequencing agriculturally important organisms, the combination of bioinformatics and functional genomics globally enhances agricultural genomics. These studies are likely to pave the way towards better understanding of plant–pathogen biological networks, and eventually to lead to breakthroughs in the promotion of plant resistance to agricultural pests.

Abbreviations: 2DE – 2-dimensional gel electrophoresis; AFLP – amplified fragment length polymorphisms; EST – expressed sequence tag; MS – mass spectrometry; UTR – untranslated region; Y2H – yeast two hybrid.

Introduction

The information age is upon us and its consequences are profound; they amount to a revolution in the natural sciences. Genome sequencing, the orderly reading of the millions of nucleotides constituting the DNA of a living organism, produces one of the major challenges of these days: that of interpreting the enormous amount of sequencing data in order to obtain useful knowledge. The practical implementation of the knowledge coming from genomics is likely to lead to breakthroughs in all areas of biology, including agricultural research.

Initially, genomics programs were focused on humans and model organisms. The human genome sequence has been completed (3000 Mb; Genome International Sequencing Consortium, 2001), as have those of the yeast *Saccharomyces cerevisiae* (12 Mb; Mewes et al., 1997), the roundworm *Caenorhabditis elegans* (97 Mb; The *C. elegans* Sequencing Consortium, 1998), *Arabidopsis thaliana* (110 Mb; Arabidopsis Genome Initiative, 2000) and the fruit fly *Drosophila melanogaster* (180 Mb; Adams et al., 2000). The complete genomes of over 600 organisms, as well as those of organisms for which sequencing is in progress, representing all the main domains of life, can be found in the World-Wide Web at the Entrez Genomes site: (http://www.ncbi.nlm.nih.gov/Entrez/Genome/main_genomes.html). Sequencing efforts are now turning to a number of agriculturally important organisms, whose elucidation will greatly facilitate the use of agricultural genomics studies to enhance our understanding of agriculture-related biology. A draft of the sequence of the rice genome has recently been released, and the target date for the final version is 2004 (Goff et al., 2002; Yu et al., 2002). Also in progress is the sequencing of the genome of the model legume *Medicago truncatula*. In addition, several major genomics and expressed sequence tag (EST) sequencing programs of plant pathogens were initiated recently (see below), while sequencing of the genome of the beneficial microorganism *Sinorhizobium meliloti* is complete (Capela et al., 2001).

Agricultural genomics may be defined as the application of genomic tools to study agriculture-related questions. Within the scope of plant pathology, these

studies typically focus on the functions of groups of genes within crops and pests, and their interactions. Traditionally, agricultural genetic research has involved a 'one gene at a time' approach, focused on understanding the influence of single genes. The results of this type of study have lead to breakthroughs in the biotechnology industry and major advances in medicine and other fields. However, a gene rarely works alone; virtually every aspect of a living creature is determined and affected by a genetic network comprising several or many genes. Agricultural genomics, which focuses on understanding a crop or a pathogen by simultaneously studying all or most of its genes (instead of a single gene at a time), is generating information about the functions of networks of genes, including those that correspond to agriculturally valuable crop traits, such as disease resistance or tolerance.

This review covers some of the recent studies, which utilized genomics approaches to examine plant pathology. Emphasize is placed on the importance and limitations of interdisciplinary studies for genome-scale data analysis. Such studies combine bioinformatics, the computational study of sequence data with functional genomics, the high-throughput examination of gene products and functions. In addition, future prospects and pitfalls for application of this novel biological knowledge to the improvement of plant protection will be discussed.

Plant pathology-related data resources and bioinformatics

Several major genomics initiatives concerning model plants, or plants of agricultural importance are in progress or they have already led to the publication of genomes. Genomics centers for agricultural crop plants include The Tomato Center (http://www.sgn.cornell.edu/), The Rice Center (http://www.tigr.org/tdb/e2k1/osa1/intro.shtml), The Legume *Medicago truncatula* Center (http://www.medicago.org/), The Banana Center (http://www.promusa.org/) and The Eucalypt Center (http://www.agrf.org.au/future_initiatives.html). In contrast to the relatively small genomes of plants that have already been sequences, or are in the process of being so, complete sequencing of plants with large genomes is not yet practicable, especially for those with genomes larger than the 3000 Mb human genome, such as barley (5000 Mb) and wheat (16,000 Mb). A more reasonable approach to gene discovery in such large-genome crop plants is through the development of databases of ESTs, which provide a wealth of information in a relatively short time. An EST is a set of single-pass sequenced cDNAs from an mRNA population derived from a specified cell population (representing, for example, a specific tissue, organ, developmental state or environmental condition).

On October 22, 2002 the EST database (dbEST, GenBank, accessible at http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html) contained 12,997,157 public entries, and it continues to grow daily. The interpretation of such huge amounts of data comprises not only a biological challenge but is also a major computational challenge. The race between database growth and proficient query capabilities continues and computational improvements are continuously needed to support efficient biological discoveries. Agriculturally relevant plants, as well as model plants used in agricultural related research, for which there are major collections of ESTs include: *Glycine max* (soybean) with 290,224 ESTs; *Hordeum vulgare* ssp. *vulgare* (barley), 240,885 ESTs; *A. thaliana* (thale cress), 176,915 ESTs; *Triticum aestivum* (wheat), 257,022 ESTs; *M. truncatula* (barrel medic), 170,500; *Zea mays* (maize), 180,919; *Lycopersicon esculentum* (tomato), 148,373; *Oryza sativa* (rice), 110,132; *Sorghum bicolor* (sorghum), 84,712 and *Solanum tuberosum* (potato), 94,420 ESTs. Among the coniferophyta, *Pinus taeda* (loblolly pine) leads with a collection of 60,226 ESTs and the anthophyta, *Populus tremula* × *Populus tremuloides* heads the list with 20,084 ESTs.

Several major genomic and EST sequencing programs of plant pathogens were initiated recently. Those are anticipated to promote pathogen genomics and greatly benefit agricultural research. The National Human Genome Research Institute recently announced a program to review proposed fungi for priority for genome sequencing. The Steering Committee of the Fungal Genome Initiative (FGI), in collaboration with the Whitehead Institute/MIT Center for Genome Research, is reviewing candidates for sequencing. These include important plant pathogens such as *Ustilago maydis* (agent of smut disease on corn) and *Fusarium graminearum* (agent of Fusarium head blight [scab] of wheat and barley) (http://www-genome.wi.mit.edu/seq/fgi/candidates.html). In addition, a genome project for *Magnaporthe grisea* (agent

of rice blast disease) was initiated as a partnership between the International Rice Blast Genome Consortium and the Whitehead Institute for Biomedical Research (http://www-genome.wi.mit.edu/annotation/fungi/magnaporthe/). Other databases of fungal plant pathogens include The Phytopathogenic Fungi and Oomycete EST Database (http://cogeme.ex.ac.uk/). Examples of plant pathogenic bacterial genome projects are the TIGR initiative for the sequencing of *Pseudomonas syringae pv. tomato* (http://www.tigr.org/tdb/mdb/ mdbinprogress.html), and the recent completion of the *Ralstonia solanacearum* 5.8 Mb genome (Salanoubat et al., 2002). Entries for other pathogens include the plant-parasitic nematodes *Meloidogyne incognita* with 12,752 ESTs and *Meloidogyne arenaria* with 3334 ESTs. The symbiotic mycorrhizal fungus, *Glomus intraradices* has 2963 ESTs.

Despite ESTs being typically short and of relatively low quality, with a majority representing only the 3' untranslated region (UTR), which makes their functional annotation difficult, ESTs are useful molecular landmarks. They provide a profile of the mRNA population and offer a quick method for cloning a large number of genes known to be expressed in a cell population. Those from groups of tissues can be used to answer questions about tissue-specific genes and to identify novel proteins (Allikmets et al., 1995; Braren et al., 1997). In particular, useful information can be found by clustering ESTs and mRNAs based on sequence overlaps, to yield sequences that are longer and more accurate, and which, therefore, better represent the underlying genes. Multiple aligment of ESTs from a given gene may reveal polymorphism, and EST databases may be used for *in silico* expression-profiling across multiple libraries.

We are currently analyzing 30 different *M. truncatula* EST libraries (accessible at http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/map00?taxid=3880) in order to find genes involved in the various plant–microbe interactions. Expression profiles are being analyzed using a modification of the 'guilt by association' method (Walker et al., 1999), annotating the sequences, dividing them into functional groups and characterizing the presence of sequences specific to certain interactions (rhizobium–plant, mycorrhiza–plant, pathogen–plant, etc.). For example, the library of *M. truncatula* roots (3156 ESTs) that had been infected with the nematode *M. incognita* (GeneBank submission authors: Bird, D., Koltai, H., Samac, D., Town, C.D., Van Aken, S., Utterback, T., Cheung, F., Tsai, J., Fraser, C.M., 2002) was subtracted from the rest of the submitted *M. truncatula* EST libraries. The entire library was annotated and, following the subtraction several classes of plant ESTs that are specific to the nematode-*M. truncatula* interaction were found (Volpin and Koltai, unpublished). Among them are transcriptional factors and protein kinases, which may be involved in the signal transduction pathways activated during the early events of nematode infection.

Plant pathology-related functional genomics

Transcriptomics and proteomics are among the powerful tools for the quantitative, real-time study of gene expression on the genome-scale. They enable us to assign products and functions to a majority of the genes that comprise an organism. Transcriptomics, which aims to profile gene transcription in a cell or a tissue globally, is mainly applied either by *in situ* synthesis of oligonucleotides ('oligonucleotide microarrays') or by deposition of pre-synthesized DNA fragments ('cDNA microarrays') on solid surfaces and their hybridization with labeled gene transcription products (reviewed by Aharoni and Vorst, 2002). The direct proteomics approach may combine 2-dimensional gel electrophoresis (2DE) of proteins with mass spectrometry (MS), to generate a catalog of expressed proteins (reviewed by Kersten et al., 2002), whereas reverse proteomics may globally map protein interactions by means of yeast two-hybrid (Y2H) system (Fang et al., 2002; Walhout and Vidal, 2001).

Plant response to pathogens

Active resistance of plants to pathogens depends on recognition of the pathogens and initiation of defense mechanisms. Despite the identification of the R genes of many plants, little is known about the signal transduction pathway that confers resistance, or about the events that occur during the interaction between a parasite and a compatible plant, and that establish and maintain parasitism. Traditionally, studies of the interactions of pathogens with plants were focused on one or only a few components of the interaction, and thus yielded only limited information on the genetic networks involved. Nevertheless, from relatively few extensive functional genomics studies, an impressively

large amount of information has been gained, on plant microbe interactions.

Using cDNA arrays comprising 13,000 unique ESTs, Scheideler et al. (2002) monitored global changes in the *A. thaliana* transcriptome (i.e. the whole set of transcripts present at a given time point in a cell, tissue or organism) after attempted infection with the incompatible bacterial pathogen *P. syringae* pv. *tomato*. A massive transient shift in gene expression from primary, housekeeping metabolism to defense-related metabolism was observed within the first 24 h after inoculation of the plant with the incompatible pathogen. Differentially expressed ESTs included those coding for metabolic enzymes, cellular organization proteins, signal transduction proteins and proteins involved in control of gene expression and stress responses (Scheideler et al., 2002). The shift in gene expression may reflect important changes in plant metabolism, some of which were not suspected previously to be associated with the plant defense response (Scheideler et al., 2002).

Conversion of genetic pathways

Microarray studies have indicated that different defensive signaling pathways converge, to form substantial networks that control and coordinate regulatory interactions. Schenk et al. (2000) exposed *A. thaliana* to the incompatible fungal pathogen *Alternaria brassicicola*, or treated it with various defense-related signaling molecules, such as salicylic acid (SA), methyl jasmonate (MJ) or ethylene. A study of 2375 ESTs that represented a biased population of putative defense-associated and regulatory genes, identified genes that were differentially expressed in response to each of the treatments (pathogen, SA, MJ and ethylene), as well as a surprisingly large group of genes that were coordinately expressed in SA, MJ and ethylene treatments (Schenk et al., 2000). Similarly, the expression of several *A. thaliana* transcription factors was reduced in all mutants in which the SA, MJ or ethylene signaling pathways were blocked (Chen et al., 2002). In contrast, some transcription factors whose expressions are reduced in mutants defective in SA signaling, exhibited enhanced expression in mutants defective in either MJ or ethylene signaling (Chen et al., 2002). Together, these results strongly suggest the existence of both positive and negative interactions among defensive signaling pathways.

Several sets of evidence suggest that genetic networks, which are involved in plant responses to pathogens and to abiotic stress, converge. Microarray studies identified a cluster of five *A. thaliana* transcription factors that were induced by both biotic and abiotic (water) stresses (Chen et al., 2002). In addition, heat-, salt-, drought-, or cold-stress inducible genes were upregulated in *A. thaliana* infected with the incompatible bacterial pathogen *P. syringae* pv. *tomato*, especially following 24 h of infection (Scheideler et al., 2002). These finding matched previous results obtained by Durrant et al. (2000), who used an RNA fingerprinting technique of cDNA amplified fragment length polymorphisms (AFLPs). Durrant et al. (2000) used tobacco cell culture expressing the tomato *Cf-9* (the gene conferring resistance in tomato to the fungus *Cladosporium fulvum*) to demonstrate that the expression of *ACRE* (for Avr9/Cf-9 rapidly elicited) genes was induced in response to either a fungal elicitor or a physical stress.

In addition, Chen et al. (2002) suggested that plant responses to different pathogens share some features; they found that some Arabidopsis transcription factors were preferentially activated in response to a bacterium (*P. syringae*), a fungus (*Botrytis cinerea*) and a virus (*Cauliflower mosaic virus*).

Epistatic interactions between components of signal transduction pathways

A direct proteomics approach that combined phospholabeling, 2DE and nanoelectrospray ionization tandem MS (nanoESI-MS-MS) was used by Peck et al. (2001) to identify proteins that were rapidly phosphorylated during the response of Arabidopsis cells to bacterial flagellin. A number of proteins showed decreased or increased levels of phosphorylation within minutes of flagellin treatment; among them was the protein AtPhos43, whose phosphorylation was dependent on FLS2, a receptor-like kinase involved in flagellin perception. This suggests that AtPhos43 is a component of the signal transduction pathway of pathogen perception, downstream to the signal receptor (Peck et al., 2001).

Protein interactions between genetic pathway components

Fang et al. (2002) used reverse proteomics to examine protein–protein interactions. Proteins expressed in yeast from a cDNA library of 7-day-old rice seedlings, which had been inoculated with spores of the rice blast fungus *M. grisea*, were used as prey. Various rice proteins related to resistance and defense

signal-transduction pathways were used as baits. They included the rice homolog to the Arabidopsis NPR1, a key regulator of defense responses in Arabidopsis. Fang et al. (2002) found that NPR1 interacts not only with bZIP transcription factors, as previously suggested (Zhang et al., 1999), but also with serine-threonine receptor protein kinase, GTP-binding proteins and host defense-response proteins. This suggests that these proteins have a role in *NPR1*-mediated signaling (Fang et al., 2002).

In addition, different sets of proteins interacted with the rice homologs of Pti4/5/6 proteins (Fang et al., 2002). These ERF-family members share considerable homology in the conserved EREBP domain, and interact with Pto kinase to confer resistance to *P. syringae* in tomato (Zhou et al., 1997). Perhaps these findings indicate the discrete roles of Pti4, 5 and 6 in plant defense (Fang et al., 2002). Evidently, expression of tomato *Pti4* gene in Arabidopsis caused an increase in the steady-state abundance of PR gene transcripts, whereas transgenic plants expressing *Pti5* or *Pti6* showed only weak or no increase in the expression of those PR genes (Gu et al., 2002).

Regulation of gene expression

One of the most powerful facilities provided by high-throughput gene expression-profiling is the ability to examine the progression of regulated functions, in which different categories of transcripts show regulation at different time points. This enables *cis* regulating elements (and the involvement of *trans* regulating elements) to be identified. Several systemic acquired resistance (SAR)-responsive genes have been found to be co-expressed with *PR-1*; their common promoter contained a core binding site for WRKY transcription factor (Maleck et al., 2000). In addition, the presence of a highly conserved WRKY-like promoter motif in genes induced during infection with a broad set of pathogens – including bacteria, fungi, oomycetes and viruses – suggested that other WRKY-like factors may also be involved in pathogen-related responses (Chen et al., 2002).

WRKY was hypothesized to be a common repressor of *PR-1* regulon genes (i.e., genes with common regulation patterns), which were de-repressed by the acquisition of SAR (Maleck et al., 2000). The involvement of WRKY DNA-binding proteins in pathogen-induced signaling pathways was highlighted by the identification and analysis of the *A. thaliana RRS1* alleles involved in determining recessive resistance to several strains of the causal agent of bacterial wilt, *R. solanacearum*. Identification of a WRKY motif in RRS1-R suggested that the activation of the C-terminus WRKY motif might activate a signaling cascade, or directly activate defense-related genes, so leading to the plant resistance response (Deslandes et al., 2002).

Arabidopsis as a model for plant–pathogen interaction

The wide range of available bioinformation on Arabidopsis makes it a good facilitator for understanding the functions of genes and proteins. Nevertheless, despite the conservation of the gene repertoire between Arabidopsis and other *Brassica* species (Paterson et al., 2001), there is only limited co-linearity (microsynteny), over small chromosomal segments, between Arabidopsis and tomato (reviewed by Mysore et al., 2001). This suggests that the establishment of synteny between genomes of species belonging to families as divergent as those of Arabidopsis and tomato may be difficult. Despite the high proportion of the predicted Arabidopsis proteins that are significantly homologous to those of rice (85%; draft of rice genome sequence released April, 2002; Goff et al., 2002; Yu et al., 2002), only 2% of the syntenic protein pairs (two proteins found in close proximity in both rice and Arabidopsis) on Arabidopsis chromosome 5 are adjacent to one another; most are separated by 1–150 intervening proteins (Goff et al., 2002). In addition, only limited co-linearity was found between the Arabidopsis and maize genomes, and a significant proportion of the maize ESTs encode highly diverged or maize-specific proteins (Brendel et al., 2002). To complicate the picture even further, many evolutionary orthologous genes do not show identical expression patterns between different plant species and although having the same biochemical function, do not seem to be related to the same biological processes (Volpin unpublished).

Despite the limited similarities in genome organization and function between agricultural crops and Arabidopsis, most of the high-throughput studies of plant pathology have used Arabidopsis. Therefore, it is important, in general, to determine how good Arabidopsis is as a model for crops and, in particular, to decipher the rapidly evolving, highly variable mechanisms of plant–pest interactions. Many Arabidopsis R gene homologues have been identified in rice. Some, abundant in Arabidopsis, were found to be absent from rice, e.g., the nucleotide-binding, leucine-reach repeats (NB-LRRs) proteins that encode the TIR (Toll-interleukin 1 receptor) motif at their amino termini.

Others, especially those controlling disease-resistance signal transduction cascades, are redundant in rice, which prevents the assignment of putative Arabidopsis orthologues (Goff et al., 2002). In addition, in maize, the proteins that are distinct and do not have close Arabidopsis homologues include those involved in maize pathogen-defense responses (Brendel et al., 2002).

Thus, the task of correlating genetic complexes that include both susceptibility to and resistance to pathogens may prove to be much more complicated when comparison between species is involved. It is essential that studies specific to particular crop/pathogen combinations are conducted to facilitate the determination of crop susceptibility, rather than relying on models based on comparative biology.

Pathogen virulence

The global, genomics approach makes it possible to explore pathogen virulence mechanisms. Kekarainen et al. (2002) used a genome-wide, transposition-based *in vitro* mutagenesis in the functional screening of viral propagation-related genomics sites in *Potato Virus A*. Over 300 sites critical for virus propagation were identified. Many were in new locations, which had previously not been assigned to any viral function (Kekarainen et al., 2002). Also, a recent study by Guttman et al. (2002) identified 'effector' proteins secreted by a type III apparatus, which may play a central role in bacterial virulence. This study combined bioinformatics predictions based on the known genome sequence of *P. syringae* and on the amino-acid composition and conserved elements in effectors, with functional tests of the hypersensitive response (HR) used as an *in vivo* marker in Arabidopsis. Importantly, the secretion of two putative effectors that were predicted by means of bioinformatics tools was shown to be type III-dependent (Guttman et al., 2002). This illustrates the advantage of an interdisciplinary approach, combining bioinformatics and functional genomics, for the study of pathogens.

Future prospects: Genomics for the benefit of plant protection

Genomics and post-genomics studies are expected to provide an unprecedented enhancement of our understanding of the biological events in plants in general, and of plant–pathogen interactions in particular. Additional post-genomics approaches for plant pathology studies could and should be applied. They could include a reverse genetics approach involving genome-scale knockout mutations and, especially, T-DNA insertion mutants that have become a valuable resource for the study of gene function in Arabidopsis (Tax and Vernon, 2001; Thorneycroft et al., 2001). Proteomics studies using protein microarrays for the detection of immobilized antigen with antibodies or vice versa (reviewed by Kersten et al., 2002), and metabolomics studies, aiming to examine the set of metabolites synthesized by a biological system (reviewed by Fiehn, 2002) should be integrated.

Integrated genomics studies are likely to reinforce the ability not only to select the candidate genes most suitable for manipulation, in order to enhance resistance or to reduce susceptibility, but also to provide better, more accurate prediction of the outcome of such manipulation. The genetic pathways that characterize a 'plant' in its 'host' capacity, will be understood. Therefore, the influence of manipulation of one or more components of a plant's pathogen-response pathway, on other pathways that converge with the manipulated one, may be better predicted.

Similarly, any genomics approaches that lead to a global understanding of pathogens as complex organisms and highlight their agriculturally important attributes in the context of the understanding of their biology, could facilitate the design of new ways to control these organisms. For example, new herbicides and fungicides might be developed by targeting pathogen-specific genetic pathways. Conventional approaches to the development of new herbicidal and fungicidal products has involved spraying various chemicals on weeds and fungi, in the hope of finding a chemical that would affect the target without killing the crops. Once a promising chemical has been identified, labor-intensive genetic, physiological and biochemical techniques are used to identify the protein in the weed or fungus that is affected by it. This conventional approach is expensive and slow, and has a low success rate. The development of environmentally friendly herbicides and fungicides will be accelerated by the close focusing of research efforts on the discovery of chemicals that disable specific genes and pathways within weeds or fungi.

Concluding remarks

We have presented some of the recent studies of plant pathology that rely on genomics approaches. Using the growing collection of available sequence data,

and combining bioinformatics and functional genomics should lead to a greater understanding of the genetic networks that are activated during plant responses to pathogens, and those involved in pathogen virulence. Nevertheless, genomics analysis presents some major inherent difficulties, including large sets of data of variable quality, and the lack of standardization. For example, some EST libraries are redundant, some are not, and a gene may refer to either the genomic DNA that encodes a protein, to the resulting mRNA transcript, or only to the coding sequence, etc. In addition, the computational tools for the conversion of the data to usable knowledge (data mining) are still immature. This is mainly because the high-throughput approach to data collection is not aimed at a specific biological question, so that the mining becomes a significant statistical challenge.

Nevertheless, genomics is rapidly evolving. Interdisciplinary research is constantly progressing towards the improvement of computational tools and high-throughput laboratory techniques for accurate data production and analysis. The accumulation of knowledge through correct interpretation of the data, and, importantly, the correlation of the knowledge gained from many studies, both in genomics and in other fields, will not only lead to the identification of the key genes that confer important traits. It will also pave the way towards the mapping of all genetic pathways, their interactions, divergences and convergences.

It is important to keep in mind that the picture obtained by using genomics and bioinformatics approaches is an overview that is blurred by data-interpretation problems. The temporal and spatial cascades of events that take place during, for example, plant responses to pathogens, are largely ignored, and high-throughput data concerning cell-specific reaction to microbes are non-existent. In addition, statistical similarity between plant-microbe interface patterns does not necessarily reflect a convergence of biological events. Hence, genomics data and the knowledge derived from them obviously do not provide the resolution obtained by the traditional one gene at a time approach.

We are still convinced that elucidation of the larger picture, even though it is blurred, by genomics means, will facilitate the selection of candidates for the in-depth study of individual genes, proteins or processes. Thus, genomics cannot and should not be used as a stand-alone tool. Rather, it should be integrated with other research approaches, which will thoroughly study individual genes in the course of in-depth studies of biological systems. The new, budding field of agricultural genomics will pave the way towards a global understanding of plant and pathogen biology, and agriculture is likely to benefit from its application.

Acknowledgement

We would like to thank Doron Holland for critical reading of the manuscript.

References

Adams MD et al. (2000) The genome sequence of *Drosophila melanogaster*. Science 287: 2185–2195

Aharoni A and Vorst O (2002) DNA microarrays for functional plant genomics. Plant Molecular Biology 48: 99–118

Allikmets R, Gerrard B, Glavac D, Ravnik-Glavac M, Jenkins NA, Gilbert DJ, Copeland NG, Modi W and Dean M (1995) Characterization and mapping of three new mammalian ATP-binding transporter genes from an EST database. Mammalian Genome 6: 114–117

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

Braren R, Firner K, Balasubramanian S, Bazan F, Thiele HG, Haag F and Koch-Nolte F (1997) Use of the EST database resource to identify and clone novel mono(ADP-ribosyl) transferase gene family members. Advances in Experimental Medicine and Biology 419: 163

Brendel V, Kurtz S and Walbot V (2002) Comparative genomics of Arabidopsis and maize: Prospects and limitations. Genome Biology 3: REVIEWS 1005

C. elegans Sequencing consortium (1998) genome sequence of the nematode *C. elegans*: A platform for investigating biology. Science 282: 2012

Capela D et al. (2001) Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proceedings of the National Academy of Sciences USA 98: 9877–9882

Chen W et al. (2002) Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell 14: 559–574

Deslandes L, Olivier J, Theulières F, Hirsch J, Feng D-X, Bittner-Eddy P, Beynon J and Marco Y (2002) Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by the recessive *RRS1-R* gene, a member of a novel family of resistance genes. Proceedings of the National Academy of Sciences USA 99: 2404–2409

Durrant WE, Rowland O, Piedras P, Hammond-Kosack KE and Jones JD (2000) cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. Plant Cell 12: 963–977

Fang Y, Macool DJ, Xue Z, Heppard EP, Hainey CF, Tingey SV and Miao GH (2002) Development of a high-throughput yeast

two-hybrid screening system to study protein–protein interactions in plants. Molecular Genetics and Genomics 267: 142–153

Fiehn O (2002) Metabolomics – the link between genotypes and phenotypes. Plant Molecular Biology 48: 155–171

Genome International Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921

Goff SA et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296: 92–100

Gu YQ, Wildermuth MC, Chakravarthy S, Loh YT, Yang C, He X, Han Y and Martin GB (2002) Tomato transcription factors pti4, pti5, and pti6 activate defense responses when expressed in Arabidopsis. Plant Cell 14: 817–831

Guttman DS, Vinatzer BA, Sarkar SF, Ranall MV, Kettler G and Greenberg JT (2002) A functional screen for the type III (Hrp) secretome of the plant pathogen *Pseudomonas syringae*. Science 295: 1722–1726

Kekarainen T, Savilahti H and Valkonen JP (2002) Functional genomics on potato virus A: Virus genome-wide map of sites essential for virus propagation. Genome Research 12: 584–594

Kersten B, Burkle L, Kuhn EJ, Giavalisco P, Konthur Z, Lueking A, Walter G, Eickhoff H and Schneider U (2002) Large-scale plant proteomics. Plant Molecular Biology 48: 133–141

Maleck K, Levine A, Eulgem T, Morgan A, Schmid J, Lawton KA, Dangl JL and Dietrich RA (2000) The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. Nature Genetics 26: 403–410

Mewes HW, Albermann K, Bahr M, Frishman D, Gleissner A, Hani J, Heumann K, Kleine K, Maierl A, Oliver SG, Pfeiffer F and Zollner A (1997) Overview of the yeast genome. Nature 387: 7–65

Mysore KS, Tuori RP and Martin GB (2001) *Arabidopsis* genome sequence as a tool for functional genomics in tomato. Genome Biology 2: REVIEWS 1003

Paterson AH, Lan TH, Amasino R, Osborn TC and Quiros C (2001) Brassica genomics: A complement to, and early beneficiary of, the Arabidopsis sequence. Genome Biology 2: REVIEWS 1011.

Peck SC, Nuhse TS, Hess D, Iglesias A, Meins F and Boller T (2001) Directed proteomics identifies a plant-specific protein rapidly phosphorylated in response to bacterial and fungal elicitors. Plant Cell 13: 1467–1475

Salanoubat M et al. (2002) Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature 415: 497–502

Scheideler M, Schlaich NL, Fellenberg K, Beissbarth T, Hauser NC, Vingron M, Slusarenko AJ and Hoheisel JD (2002) Monitoring the switch from housekeeping to pathogen defense metabolism in Arabisopsis thaliana using cDNA arrays. Journal of Biological Chemistry 277: 10555–10561

Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC and Manners JM (2000) Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. Proceedings of the National Academy of Sciences USA 97: 11655–11660

Tax FE and Vernon DM (2001) T-DNA-associated duplication/translocations in Arabidopsis. Implications for mutant analysis and functional genomics. Plant Physiology 126: 1527–1538

Thorneycroft D, Sherson SM and Smith SM (2001) Using gene knockouts to investigate plant metabolism. Journal of Experimental Botany 52: 1593–1601

Walhout AJM and Vidal M (2001) High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. Methods 24: 297–306

Walker MG, Volkmuth W, Sprinzak E, Hodgsdon D and Klinger T (1999) Prediction of gene function by genome-scale expression analysis: Prostate cancer-associated genes. Genome Research 9: 1198–1203

Yu J et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296: 79–92

Zhang Y, Fan W, Kinkema M, Li X and Dong X (1999) Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the PR-1 gene. Proceedings of the National Academy of Sciences USA 96: 6523–6528

Zhou J, Tang X and Martin GB (1997) The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a cis-element of pathogenesis-related genes. EMBO Journal 16: 3207–3218