The Plant Cell, Vol. 13, 449, March 2001, www.plantcell.org © 2001 American Society of Plant Physiologists

IN THIS ISSUE

A Brief Tour of the Cell Cycle

The mitotic cell cycle consists of alternating rounds of DNA replication (which occurs during the S phase) and chromosome segregation (mitosis or M phase) interrupted by gaps known as G1 (the interval before S phase) and G2 (the interval after S phase). Events that occur in each phase are regulated to ensure that DNA is replicated only once in each cell cycle and that each daughter cell ends up with a complete complement of the genome. Regulation of phase transitions is principally achieved through the action of a host of cyclin-dependent kinases (CDKs) and their corresponding regulatory cyclin proteins. Although the basic cell cycle machinery is highly conserved among all eukaryotes, there are a number of important differences in cell cycle control between higher plants and other eukaryotes. In this issue of *The Plant Cell*, we highlight three articles related to different aspects of the cell cycle that focus attention on unique characteristics of plant cell cycles and point the way to new discoveries in cell cycle control and organelle replication.

Mitosis usually progresses into cytokinesis, resulting in cell division and the production of two new daughter cells. Our first stop on the cell cycle tour is to view an exception to this rule in plant cells: the development of syncytia in endosperm tissue, which is the focus of work presented by **Boisnard-Lorig et al. (pages 495–509)**. Syncytia, which are typical of plant endosperm tissue, are cells with multiple nuclei that originate from the fusion of multiple cells or from single cells in which DNA replication and nuclear divisions proceed without cytokinesis. Unlike other eukaryotes, flowering plants typically undergo double fertilization: the pollen tube delivers two male gametes to the embryo sac; one fuses with the egg to form the zygote, and the other fuses with the two polar nuclei of the diploid central cell to produce a triploid endosperm nucleus. Division of the endosperm nucleus gives rise to endosperm tissue, which surrounds the developing embryo and plays an important role in the nutrition of the embryo during embryogenesis and seed germination. In numerous species, the endosperm initially develops as a syncytium that contains up to a few hundred nuclei. Ultimately, the syncytium divides into individual cells in a process called cellularization. Boisnard-Lorig observed mitosis in the endosperm tissue of developing Arabidopsis seedlings via constitutive expression of a histone 2B::YFP fusion chromatin marker. The authors observed the development of three spatially distinct domains with syncytial endosperm that exhibit differences in the activity of cell cycle control parameters. Thus, spatial regulation of cell cycle control genes may play a role in the development of the endosperm. For example, it is shown that endoreduplication is likely limited to one of the three endosperm domains, the chalazal endosperm, because this domain accumulates very few nuclei and these nuclei are larger and appear to contain a much greater quantity of chromatin than the nuclei present in the other two domains. Endoreduplication is the most common mechanism of polyploidization in plants, a widespread phenomenon among higher plants and of particular importance in many crop species. Mechanisms of endoreduplication in plants are still poorly understood (Joubès and Chevalier, 2000), and the work of Boisnard-Lorig et al. may point the way to further knowledge of this process.

The second stop on our tour is at a spindle assembly checkpoint in the M phase of embryo cells of the brown alga Fucus. **Corellou et al. (pages 585–598)** studied aspects of the embryonic cell cycle in *Fucus spiralis*, which produces large populations of synchronously developing external zygotes that are easy to manipulate and observe experimentally. In animals, there are major differences in cell cycle control between somatic cells and embryo cells. The somatic cell cycle is tightly regulated via a series of "checkpoints," typically involving CDKs, which monitor cell cycle progression, ensuring that only one round of DNA replication occurs per cycle and preventing mitosis until DNA synthesis and repair are complete. In contrast, checkpoints involving tight regulation of CDK activity appear to be less stringent or absent from early embryonic cells in animals. From experiments using the drugs nocodazole (which inhibits mitotic spindle formation) and olomoucine (a specific inhibitor of CDK activity), Corellou et al. show that, similar to somatic cell cycles, the Fucus zygote cell cycle includes a spindle assembly checkpoint targeted at CDK activity. This checkpoint appears to operate by maintaining high levels of CDK activity, and the authors suggest that inactivation of CDKs may be required for cells to undergo chromatin decondensation and exit mitosis. Previous work by this group (Corellou et al., 2000) showed that Fucus zygotes also contain a DNA replication checkpoint. In the present work, the authors show that the initial embryonic cell cycles of Fucus resemble the somatic rather than the early embryo cell cycles of animals in a number of other important characteristics as well. For example, animal cells in early embryogenesis undergo rapid cell cycles that consist only of S and M phases. Corellou et al. show that, as in

IN THIS ISSUE

somatic cells, the Fucus embryonic cell cycle consists of well-defined G1, S, G2, and M phases. Furthermore, CDKs show complex regulation involving translational control and post-translational control via tyrosine phosphorylation. Complex regulation of CDKs is another characteristic of somatic cells that is lacking in animal embryo cells, in which CDK control over the switch between the S and M phases appears to be regulated principally by periodic synthesis and degradation of cyclin B, a positive regulator of CDK activity.

Although it is tempting to view the Fucus system as a model for photosynthetic organisms, including higher plants, it must be remembered that brown algae (phylum Phaeophyta) belong to the division of Stramenopiles (also called Chromista), a group of protists that includes the diatoms (Bacillariophyta), yellow-green algae (Xanthophyta), and water molds (Oomycota). Phaeophytes and other stramenopiles, although traditionally classified as plants, actually are not closely related to green plants and occupy their own division within the kingdom Protista (or perhaps they may be placed in a separate kingdom called Chromista). Eukaryotic phylogenies show that animals and fungi may be more closely related to plants than are brown algae and other stramenopiles. Thus, work with Fucus may not provide information about higher plant embryogenesis and must be viewed with caution in this context. Rather, this work provides valuable information for comparative studies of eukaryotes, including plants, animals, and fungi. It remains to be seen whether characteristics of cell cycle control in Fucus are shared with plants and/or other eukaryotes.

In one of the most detailed analyses of the embryonic cell cycle of higher plants to date, Sauter et al. (1998) presented reverse transcriptase–mediated polymerase chain reaction analysis of the expression of three cyclin genes and the CDK *cdc2* in sperm, egg cells, and other cells of the embryo sac isolated from maize plants and in developing zygotes produced via in vitro fertilization. Whereas *cdc2* transcripts were expressed in all gametic cells and in the zygote throughout development to the two-cell stage at 48 hr after fertilization, the cyclin genes exhibited cell-specific expression in the embryo sac and differential expression during zygote development. All cyclins were transcribed de novo after fertilization, with a high degree of expressional regulation during the first embryonic cell cycle. This is in marked contrast to findings in most animal cells, in which the egg provides a store of mRNAs, including cyclins, which control regulation of the initial embryonic cell cycles. Sauter et al. (1998) hypothesized that this zygotic regulation of the first cell cycle may contribute to a greater adaptive ability of plants during early embryogenesis, unlike animal embryos, in which patterns of cell division are more rigidly fixed before fertilization. This analysis also might lead to the prediction that cell cycle control in higher plant zygotes will show some degree of similarity to that in the Fucus model, because Corellou et al. have shown that Fucus also exhibits a high degree of zygotic regulation of the cell cycle during early embryogenesis.

The final stop on this brief tour of the cell cycle is to investigate another feature unique to plant (and algal) cells: plastids and plastid division. Plastids are believed to have originated from an endosymbiosis event in which an early photosynthetic prokaryote invaded or was engulfed by a primitive eukaryotic host (Margulis, 1970; Gray, 1992). Plants contain a number of different types of plastids in addition to chloroplasts: amyloplasts, which accumulate in seeds and tubes; leucoplasts, which are found principally in petals; and chromoplasts, which are found in fruits and flowers. All plastids differentiate from small, colorless proplastids, which are found in dividing cells in meristematic tissue. However, division of mature plastids also occurs in developing cells, and dumbbell-shaped dividing chloroplasts have been observed in young leaves of a number of species (Pyke, 1999).

Plastid division appears to be driven by the formation of a contractile ring called the plastid-dividing (PD) ring (reviewed by Pyke, 1999). A similar dividing ring controls bacterial cell division, a major component of which is FtsZ, a filamentous protein with structural similarity to tubulin. A family of *FtsZ* homologs was found in Arabidopsis, and transgenic plants constitutively expressing antisense *AtFtsZ* constructs have mesophyll cells with fewer enlarged chloroplasts compared with wild-type plants, suggesting that FtsZ proteins form part of the PD ring in plant cells (Osteryoung et al., 1998). The PD ring is a double or triple ring. Osteryoung et al. (1998) identified chloroplast-targeted and nontargeted forms of FtsZ in Arabidopsis and hypothesized that these forms make up part of the inner and outer PD rings, respectively.

Miyagishima et al. (pages 707–721) present high resolution ultrastructural analysis of the outer PD ring in dividing chloroplasts of the red alga *Cyanidioschyzon merolae*. Negative staining of isolated dividing chloroplasts showed that the outer ring consisted of 5-nm-diameter filaments, which formed a rigid structure that did not disassemble in 2 M urea . Immunoblotting of extracted proteins with CmFtsZ-specific antibody suggested that FtsZ was not present in the outer ring, and an unidentified 56-kD protein was isolated as a candidate ring component. The authors previously identified two types of FtsZ protein in *C. merolae*, one (CmFtsZ2) with similarity to cyanobacterial FtsZ that localized to chloroplast protein fractions and one (CmFtsZ1) that resembles an α-proteobacterial counterpart and is likely to function in mitochondria (Takahara et al., 2000). Antibody specific to CmFtsZ2 was used

in the present work. The authors point out the possibility that another FtsZ protein, which does not cross-react with the antibody used, forms part of the outer ring, but their results suggest that the insoluble 56-kD protein is the principal ring component, and insoluble properties have not been reported previously for FtsZ proteins.

However, as in the case of Fucus, it cannot be assumed that green plants necessarily will share features of red algae, which belong to the protist phylum Rhodophyta. In transgenic Arabidopsis plants, expression of either *AtFtsZ1* or *AtFtsZ2*, both of which encode proteins related to cyanobacterial FtsZ sequences, reduces the number of chloroplasts in mature leaf cells from 100 to 1, indicating that both genes are required for chloroplast division in higher plant cells (Osteryoung et al., 1998). It is reasonable to hypothesize that FtsZ2 constitutes part of the outer PD ring, because, unlike FtsZ1, it lacks a chloroplast transit peptide and is not transported into isolated chloroplasts.

The cell cycle of higher plants contains unique features, aspects of which may be ripe for discovery in Arabidopsis. For example, molecular identification of the 56-kD protein component of the outer PD ring in *C. merolae* will lead to the immediate identification of homologs in the completely sequenced genome of Arabidopsis. It also should be possible to identify all of the CDKs in Arabidopsis and begin a process of systematically investigating their possible functions in cell cycle control and other plant growth processes. The three articles highlighted here should stimulate renewed interest in comparative studies of cell cycle control among higher plants, algae, and other eukaryotes.

Nancy A. Eckardt
News and Reviews Editor

REFERENCES

Boisnard-Lorig, C., Colon-Carmona, A., Bauch, M., Hodge, S., Doerner, P., Bancharel, E., Dumas, C., Haseloff, J., and Berger, F. (2001). Dynamic analyses of the expression of the HISTONE::YFP fusion protein in Arabidopsis show that syncytial endosperm is divided in mitotic domains. Plant Cell **13,** 495–509.

Corellou, F., Bisgrove, S.R., Kropf, D.L., Meijer, L., Kloareg, B., and Bouget, F.-Y. (2000). A S/M DNA replication checkpoint prevents nuclear and cytoplasmic events of cell division including centrosomal axis alignment and inhibits activation of cyclin-dependent kinase-like proteins in fucoid zygotes. Development **127,** 1651–1660.

Corellou, F., Brownlee, C., Detivaud, L., Kloareg, B., and Bouget, F.-Y. (2001). Cell cycle in the Fucus zygote parallels a somatic cell cycle but displays a unique translational regulation of cyclin-dependent kinases. Plant Cell **13,** 585–598.

Gray, M.W. (1992). The endosymbiont hypothesis revisited. Int. Rev. Cytol. **141,** 233–357.

Joubès, J., and Chevalier, C. (2000). Endoreduplication in higher plants. Plant Mol. Biol. **43,** 735–745.

Margulis, L. (1970). Origin of Eukaryotic Cells. (New Haven, CT: Yale University Press).

Miyagishima, S., Takahara, M., and Kuroiwa, T. (2001). Novel filaments 5 nm in diameter constitute the cytosolic ring of the plastid division apparatus. Plant Cell **13,** 707–721.

Osteryoung, K.W., Stokes, K.D., Rutherford, S.M., Percival, A.L., and Lee, W.Y. (1998). Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial *ftsZ*. Plant Cell **10,** 1991–2004.

Pyke, K.A. (1999). Plastid division and development. Plant Cell **11,** 549–556.

Sauter, M., von Wiegen, P., and Lörz, H. (1998). Cell cycle regulatory genes from maize are differentially controlled during fertilization and first embryonic cell division. Sex. Plant Reprod. **11,** 41–48.

Takahara, M., Takahashi, H., Matsunaga, S., Miyagishima, S., Sakai, A., Kawano, S., and Kuroiwa, T. (2000). A putative mitochondrial *ftsZ* gene is encoded in the unicellular primitive red alga *Cyanidioschyzon merolae*. Mol. Gen. Genet. **264,** 452–460.

Agricultural Microbes Genome 2: First Glimpses into the Genomes of Plant-Associated Microbes

These are extraordinary times for the biological sciences. With the advent of genomics, the entire hereditary blueprint of an organism can be determined, and a deep transition into genome-enabled biology is occurring. This new era of research allows a comprehensive study of gene function that has already delivered tremendous insights into the basic biology of living organisms. Even though microbiologists were at the forefront of this paradigm shift, plant microbiologists had to be patient and for years watched with envy (and at times despair)

MEETING REPORT

the ever-growing list of genome sequences of model and animal-associated microbes. Due to, among other things, an unfavorable funding climate, the first completed genome of a plant-associated microbe, that of the citrus pathogen *Xylella fastidiosa*, was reported last year by a Brazilian team (Simpson et al., 2000) five years after the completion of the genome sequence of a human pathogenic bacteria (Fleischmann et al., 1995).

Fortunately, the tide has turned. Plant microbiologists attending the 2nd Agricultural Microbes Genome Conference, held from January 17 to 19, 2001, in San Diego, were greeted not only by the warm and sunny climate of southern California but also by the revitalizing news of the exponential accumulation of genome sequence data for plant-associated microbes. The gathering, which followed the popular Plant and Animal Genome Conference, was put together by an organizing committee chaired by Peter Johnson (U.S. Department of Agriculture [USDA]; Washington, DC) and Stephen Heller (USDA; Beltsville, MD) and was attended by more than 120 participants. The program, which consisted of ~25 talks and 25 posters, was focused essentially on the genomics of microbes relevant to various aspects of agricultural sciences and also included a bioinformatics workshop.

In this report, we highlight selected talks and posters on plant-associated microbes presented at the 2nd Agricultural Microbes Genome Conference. Meeting abstracts are archived on the World Wide Web at http://www.intl-pag.org/agm. A list of resources on the genomics of plant-associated microbes is shown in Table 1.

PLANT PATHOGENIC BACTERIA

Christian Boucher (Institut National de la Recherche Agronomique–Centre National de la Recherche Scientifique, Castanet Tolosan, France) discussed the virtually complete genome sequence of *Ralstonia solanacearum*, a wide host range bacterial pathogen that causes the bacterial wilt disease in more than 200 plant species from more than 40 botanical families, which was obtained in collaboration with the French genomics center Genoscope. The selected strain, GMI1000, infects several solanaceous crops, such as potato, tomato, and eggplant, as well as the model plant *Arabidopsis thaliana*, and induces a hypersensitive response on tobacco. The *R. solanacearum* GMI1000 genome consists of two replicons, a

Table 1. Internet Resources on Genomics of Plant-Associated Microbes

Taxonomic Group/Microbe Species[a]	URL[b]
Bacteria; Proteobacteria: Alpha subdivision; Rhizobiaceae group	
Sinorhizobium meliloti	http://sequence.toulouse.inra.fr/meliloti.html
Bradyrhizobium japonicum	http://www.genome.clemson.edu/~twood/projects/brady.html
Mesorhizobium loti	http://www.kazusa.or.jp/rhizobase
Rhizobium sp strain NGR234	http://genome.imb-jena.de/other/cfreiber/pNGR234a2.html
Beta subdivision; Ralstonia group	
Ralstonia solanacearum strain GMI1000	http://sequence.toulouse.inra.fr/R.solanacearum
Gamma subdivision; Xanthomonas group	
Xanthomonas axonopodis pv *citri*	http://genoma4.iq.usp.br/xanthomonas
Xylella fastidiosa Citrus Variegated Chlorosis strain	http://onsona.lbi.ic.unicamp.br/xf
X. fastidiosa Pierce's disease Temecula strain	http://onsona.lbi.ic.unicamp.br/xf-grape
X. fastidiosa almond/oleander strains	http://spider.jgi-psf.org/JGI_microbial/html
Gamma subdivision; Pseudomonadaceae	
Pseudomonas syringae pv *tomato* DC3000	http://www.tigr.org/tdb/mdb
Bacteria; Firmicutes: Bacillus/Staphylococcus group	
Bacillus subtilis strain 168	http://genolist.pasteur.fr/SubtiList
Mollicutes	
Spiroplasma spp	http://www.oardc.ohio-state.edu/spiroplasma
Eukaryota: Fungi; Ascomycota	
Magnaporthe grisea	http://www.cals.ncsu.edu/fungal_genomics
Stramenopiles; Oomycetes	
Phytophthora spp	http://www.ncgr.org/research/pgi

[a] Phylogenetic subdivisons follow NCBI Taxonomy Database (http://www.ncbi.nlm.nih.gov/Taxonomy).
[b] One representative URL is provided for each organism based essentially on information obtained at the conference.

MEETING REPORT

chromosome of 3.6 Mb that bears most of the housekeeping genes and a megaplasmid of 2.1 Mb that essentially contains genes encoding adaptive functions. *R. solanacearum* is known to develop a natural competence state during infection (Bertolla et al., 1999), and the genome can be expected to contain numerous sequences acquired through lateral gene transfer. Indeed, several regions, ranging in size from 1 to 30 kb, exhibit aberrant G + C contents and are associated with sequences similar to those of genetically mobile elements.

Boucher and collaborators also used a series of criteria, such as similarity to known virulence factors, linkage to candidate virulence genes, promoter motifs, and sequence motifs characteristic of eukaryotic proteins, to identify ~150 novel candidate pathogenicity genes. Some of these candidates encode homologs of type IV pili proteins, nonfibrillar adhesins, and hemolysin that may suggest previously unknown mechanisms of interaction between *R. solanacearum* and its hosts. Remarkably, nine genes, related to the previously characterized *popC* (Gueneron et al., 2000), showed significant similarity to the eukaryotic family of leucine-rich repeat plant disease resistance genes. In addition, ~50 genes harbored promoters with PIP box motifs, which are thought to play a role in the transcriptional activation of virulence genes (Fenselau and Bonas, 1995; Zhu et al., 2000). Knockout mutants of the candidate genes are now being generated. Mutagenesis experiments with one of the candidate virulence genes, *PopB*, indicated that this gene mediates avirulence of *R. solanacearum* GMI1000 in petunia and virulence in Arabidopsis.

The Gram-negative Xanthomonas species infect a variety of crop plants, including rice, cotton, citrus, grape, and bean. In the wake of the success of the Xylella genome project, a network of laboratories from São Paulo State, Brazil, embarked on the genomic analysis of xanthomonads. Fernando Reinach (Universidade de São Paulo, Brazil) discussed the complete genome sequence of *Xanthomonas axononpodis* pv *citri* (formerly known as *Xanthomonas citri* and the cause of the economically important citrus canker disease) and the 8X coverage genome sequence of *Xanthomonas campestris* pv *campestris*, a well-studied pathogen of crucifers. *X. a. citri* harbors a 5,175,422-bp chromosome with a large number of mobile elements and ~300 scattered regions containing putative pathogenicity genes. Analyses of the 4266 open reading frames (ORFs) revealed a large number of genes encoding degradative enzymes that may function in parasitism. For example, 70 genes encoding different classes of extracellular proteases were identified. In addition to the chromosome, two interdependent low copy number plasmids, pXa33 (33 kb) and pXa65 (65 kb), also occur in *X. a. citri*. Each of these plasmids contains two copies of *pthA*, a multifunctional gene characterized by 102-bp tandem repeats and previously demonstrated to mediate host specificity, pathogenicity, and avirulence (Swarup et al., 1992). pXa65 also was found to contain a homolog of the avirulence gene *AvrPphE* from the bean pathogen *Pseudomonas syringae* pv *phaseolicola*.

Preliminary comparative genomic analyses of the two Xanthomonas species revealed a number of insertions and/or deletions and noncolinear regions, suggesting significant genome fluidity in xanthomonads. Notable differences also were observed between the *X. a. citri* genome and that of the phylogenetically related *Xylella fastidiosa*. In particular, the *X. a. citri* genome bears a larger number of genes involved in macromolecule degradation and has fewer integrated phage sequences. Clearly, these new resources should prove useful in understanding the variety of Xanthomonas–host interactions of economic consequence.

The xylem-limited plant pathogen *Xy. fastidiosa* consists of a complex group of strains with different plant hosts that are vectored by homopteran insects, particularly xylem-feeding sharpshooter leafhoppers (Cicadellinae). The mechanism of plant infection by Xylella is poorly understood. It is unclear whether Xylella simply clogs up the xylem tubes, subsequently affecting water transport and causing starvation, or if disease is caused by other means. The genome sequence of an *Xy. fastidiosa* strain that causes citrus variegated chlorosis (CVC), a devastating disease that affects ~30% of all citrus trees in Brazil, was reported last year (Simpson et al., 2000). *Xy. fastidiosa* strain CVC harbors a 2,679,305-bp chromosome with 2782 ORFs. Analyses of the genome sequence revealed a number of candidate genes that may function in virulence. For example, several genes were identified that encode extracellular polysaccharide synthesis enzymes, such as fimbriae, afimbrial adhesins, and hemagglutinins, which may function in clumping of bacteria and adhesion to xylem tubes and sharpshooter leafhopper foreguts.

Another important disease caused by *Xy. fastidiosa* is Pierce's disease (PD) of grapevine, an emerging problem that is threatening the wine industry on the West Coast of the United States (Purcell and Saunders, 1999). Even though *Xy. fastidiosa* strains that cause PD have been known in California for many years, the disease has turned epidemic since the introduction of the glassy-winged sharpshooter (*Homalodisca coagulata*), which is capable of faster and longer distance movement into vineyards than other sharpshooters, reaches higher population numbers, and actively feeds on dormant grapevines in winter, resulting in chronic Xylella infections (see http://www.cnr.berkeley.edu/xylella). Marie-Anne van Sluys (Universidade de São Paulo, Brazil) presented an advanced sequence draft of the genome of an *Xy. fastidiosa* PD strain isolated

MEETING REPORT

recently from the Temecula wine-producing region of southern California. Similar to the *Xy. fastidiosa* CVC strain, isolate Temecula contains numerous phage-derived sequences and bears no avirulence-like genes or bacterial type III secretion system genes. However, one genomic region that contains putative virulence and hemagglutinin genes in the CVC strain appears deleted in the Temecula isolate.

Additional genomics resources for Xylella are available as well. Dan Drell (United States Department of Energy, Germantown, MD) discussed the first microbial sequencing month held in October 2000 by the Department of Energy's Joint Genome Institute (JGI). This initiative resulted in 8X drafts of the genome sequences of 15 microbes, including two *Xy. fastidiosa* strains isolated from almond and oleander (http://www.jgi.doe.gov/tempweb/JGI_microbial/html). Future comparative genomic analyses of Xylella should reveal genes involved in host specificity and pathogenesis.

Understanding the molecular basis of *Xy. fastidiosa* pathogenesis and interaction with its insect vectors has been limited by the lack of genetic tools. In a poster presentation, Magalie Guilhabert (University of California, Davis) described progress in generating transposon insertions in two *Xy. fastidiosa* PD strains after electroporation of a hyperactive Tn*5* transposase-transposon complex (transposome system; Epicenter Technologies, Madison, WI). Guilhabert detected *Xy. fastidiosa* transformants with single, independent, and stable insertion events and will use them in phenotypic analyses.

A poster presentation by Alan Collmer (Cornell University, Ithaca, NY) and collaborators introduced the objectives of a multi-institutional sequencing effort of the genome of the tomato pathogen *Pseudomonas syringae* pv *tomato* strain DC3000. This project, recently funded by the National Science Foundation Plant Genome Research Program, will also include comparative genomic analyses of pseudomonads as well as functional analyses and transcriptional profiling of *P. syringae* interaction with tomato.

Kenneth Bell (Scottish Crop Research Institute, Dundee, Scotland) presented a poster on the sequence of two bacterial artificial chromosomes (BACs) of *Erwinia carotovora* subsp *atroseptica*, an important pathogen that causes soft rot on potato. The sequence revealed many potential pathogenicity genes, such as an entire *hrp* gene cluster, pectinolytic genes, homologs of the Erwinia virulence genes *dspE*, *hacAB*, and *pecSM*, homologs of Xylella hemolysin genes, and homologs of *Agrobacterium tumefaciens* pathogenicity genes. These data suggest that *E. c. atroseptica* pathogenesis may be more complex than previously thought.

Members of the genera Spiroplasma and Phytoplasma (class Mollicutes) have been associated with more than 200 plant diseases worldwide and are dispersed by homopteran insects. Spiroplasmas and phytoplasmas are restricted to the phloem tissue of plant hosts and, interestingly, replicate in their insect vectors as well. Members of the class Mollicutes arose from a Gram-positive Clostridium-like ancestor and are characterized by a lack of cell wall, small genome size (530 to 2200 kb), and low G + C content. The genome sequences of several human and animal pathogenic mollicutes have been determined, including that of *Mycoplasma genitalium*, which at 580,070 bp represents one of the smallest genomes of a cellular organism (Fraser et al., 1995). To date, the complete genome sequence of a plant pathogenic mollicute has not been determined. A poster presentation by Saskia Hogenhout (Ohio State University, Wooster) summarized the results of sample genome sequencing of one of the three plant pathogenic spiroplasmas, the corn stunt spiroplasma (CSS; *Spiroplasma kunkelii*). Spiroplasmas are primarily insect pathogens, but they also include epiphytes such as *Spiroplasma floricola* and three plant pathogens vectored by phloem-feeding leafhoppers. Comparative genomic analyses of spiroplasmas and mycoplasmas revealed that the 1600-kb CSS genome harbors a surprisingly large number of virus-derived sequences. Moreover, the CSS genome clearly bears more genes commonly found in genomes of Gram-positive bacteria, confirming earlier findings that spiroplasmas did not experience extensive gene loss, as did most members of the mollicute class. Genes predicted to confer the spiral-shaped cell structure unique to spiroplasmas were identified as well.

Whereas spiroplasmas can be cultured, phytoplasmas cannot grow outside their insect or plant hosts. Despite the difficulty of obtaining sufficiently pure phytoplasma DNA, Lia Lieftink and Bruce Kirkpatrick (University of California, Davis) presented a poster on the successful cosmid cloning of a large proportion of the 670-kb genome of Western X disease phytoplasma, a significant pathogen of peach and cherry trees in California. This library will provide the basis for obtaining the first large sequence data set for a phytoplasma genome, which in turn will aid other phytoplasma genome sequencing initiatives. Genome comparisons among epiphytic, insect, and plant pathogenic spiroplasma species, as well as between spiroplasmas and phytoplasmas, will reveal genes involved in plant pathogenesis and insect transmission of bacteria from Gram-positive ancestry.

BENEFICIAL BACTERIA

Plant-associated bacteria are not always detrimental; they also can have beneficial effects on plants, for example, by enhancing nutrient uptake and plant growth. The alpha bacterium *Si-*

MEETING REPORT

norhizobium meliloti (formerly known as *Rhizobium meliloti*) has taken such a role to a high level of sophistication. *S. meliloti* can establish intimate symbiotic interactions with legume plants, inducing the formation of root nodules and allowing nitrogen fixation and uptake by the plant (Viprey et al., 2000). Even though *S. meliloti* has been studied intensively at the genetic level, Sharon Long (Stanford University) illustrated how the complete genome sequence revealed unexpected aspects of the biology of this bacterium. *S. meliloti* strain 1021 carries a tripartite genome consisting of a 3.65-Mb chromosome and two plasmids, pSymA (1.35 Mb) and pSymB (1.68 Mb), known as the symbiotic plasmids because they carry genes required for symbiosis. A team of European and North American scientists completed the sequence of the three replicons, and an annotation jamboree was held in Toulouse, France, in September 2000. Analyses of the relation among the three replicons suggested that the pSym plasmids form an integral part of the *S. meliloti* genome. For example, pSymB, which is known to be required for growth, turned out to carry a *tRNA-Arg* gene that is missing in the chromosome. pSymA carries numerous genes with predicted functions as diverse as denitrification, nitrate respiration, and opine catabolism. The overall analyses suggest that the pSym plasmids encode multiple functions and contribute not only to the establishment of symbiosis but also to survival in the soil and other stages of the life cycle of *S. meliloti*.

Long also discussed how the complete genome sequence of *S. meliloti* was compared with sequences from other rhizobacteria, such as *Bradyrhizobium japonicum*, *Mesorhizobium loti*, and *Rhizobium* sp strain NGR234. Unlike *Bra. japonicum* and *Mes. loti*, symbiosis genes in *S. meliloti* are dispersed and not clustered in gene islands. Several symbiosis and nitrogen fixation genes that have not been identified previously by classic genetic screens were revealed by similarity searches. For example, pSymA contains orthologs of the genes *nodV* and *nodW*, which encode a two-component regulatory system that positively regulates *nod* gene expression in *Bra. japonicum* in response to plant isoflavone signals (Loh et al., 1997).

In *Rhizobium* sp NGR234, a gene cluster encoding the components of a bacterial type III secretion system was identified recently, and mutant analyses of two of these genes showed that they are required for normal nodule induction on several legume plants (Viprey et al., 1998). Interestingly, genes for the type III secretion system were absent in *S. meliloti*, indicating that perhaps a different secretion mechanism is used in this species. Indeed, the overall message from the early comparative genomic analyses of rhizobacteria is that each species has developed unique solutions and adaptations to the symbiotic lifestyle.

The family of plant-beneficial bacteria includes an illustrious member, the spore-producing *Bacillus subtilis*, which is the best studied Gram-positive bacterium and is a classic model organism for the study of developmental processes in prokaryotes. A little known fact about *B. subtilis* is that it leads a normal life outside the laboratory in soil habitats, where it is known to attach to roots and enhance plant growth. Indeed, strains closely related to *B. subtilis* are established members of the so-called plant growth-promoting rhizobacteria (Andrews and Harris, 2000), and dozens of commercial formulations of these bacteria are available to farmers. Colin Harwood (University of Newcastle upon Tyne) discussed how his research team is taking *B. subtilis* back to its roots. It appears that *B. subtilis* strain 168, the classic laboratory strain that was used in the 1997 genome sequencing project (Kunst et al., 1997), exhibits biocontrol and plant growth enhancement in tomato, possibly through the production of fungicides and plant hormones and/or the relocation of metal ions and phosphates. Analyses of the more than 4000 ORFs of *B. subtilis* revealed a number of candidate genes that could modulate plant growth promotion. For example, a number of *B. subtilis* genes, such as *yqkF*, are related to auxin-regulated genes from tobacco and Arabidopsis and may manipulate hormonal processes in plants. Other genes show similarity to chitinases or are predicted to mineralize phosphates and may help *B. subtilis* compete with fungi and other microbes in the rhizosphere niche. Harwood and collaborators have initiated high throughput mutagenesis experiments with the objective of creating knockout mutants for all *B. subtilis* genes. Mutants in the identified candidate genes will then be tested for alteration of plant growth promotion and other phenotypes. In addition, Harwood's team has initiated signature-tagged mutagenesis experiments (Hensel et al., 1995) using a *B. subtilis* strain marked with the green fluorescent protein to identify mutants affected in various test environments, such as the rhizosphere. Recovery and identification of these mutants is expected to be accelerated significantly by the availability of the complete genome sequence.

PLANT PATHOGENIC FUNGI

Diseases caused by fungi are well established as major constraints to food and crop production. Rice blast, caused by the ascomycete fungus *Magnaporthe grisea*, is one of the most devastating diseases of crop plants. In addition to its economic importance, *M. grisea*, a relative of the model fungus *Neurospora crassa*, has emerged as a model experimental system for the genetic analysis of plant pathogenic fungi. Driven by a dynamic community

MEETING REPORT

of researchers, a number of tools for classic and molecular genetic manipulation have been developed for *M. grisea*. Ralph Dean (North Carolina State University) discussed the goals of the recently funded USDA-Initiative for Future Agriculture and Food Systems (IFAFS) project that officially launches the rice blast genome project. *M. grisea* has a 40-Mb genome divided into seven chromosomes. Dean plans to generate a 5X draft sequence for chromosome 7, which at 4.2 Mb is the smallest chromosome of *M. grisea*. A BAC-by-BAC approach will be taken, based on the availability of a complete physical map of chromosome 7 (Zhu et al., 1999). In addition, Daniel Ebbole (Texas A&M University), a co-principal investigator in the IFAFS project, will generate 35,000 expressed sequence tags (ESTs) from several in vitro stages to develop a unigene set for the rice blast pathogen and help design gene prediction algorithms for analysis of the genomic sequence. This project should prove an important stepping-stone for the completion of the entire genome sequence of *M. grisea*.

PLANT PATHOGENIC OOMYCETES

Oomycetes, such as Phytophthora, downy-mildews, and Pythium, form a unique branch of eukaryotic plant pathogens with an independent evolutionary history (Kamoun et al., 1999b). Among the oomycetes, Phytophthora species cause some of the most destructive plant diseases in the world, and are arguably the most devastating pathogens of dicot plants (Erwin and Ribeiro, 1996). For example, *Phytophthora infestans*, the Irish potato famine fungus, causes late blight, a disease that results in multibillion dollar losses in potato and tomato production, and *Phytophthora sojae* causes root and stem rot of soybean, a disease that can result in up to 12% losses in crop yields. Brett Tyler (University of California, Davis) introduced another recently funded USDA-IFAFS project that significantly extends the efforts of a community-based initiative known as the Phytophthora Genome Initiative (www.ncgr.org/research/pgi/whitepaper.html). Building on the successful completion of pilot cDNA sequencing projects (Kamoun et al., 1999a; Qutob et al., 2000), Tyler and collaborators plan to generate an additional 50,000 ESTs from a series of in vitro and infection stages of *P. infestans* and *P. sojae*. These sequences will be compiled in an expanded and improved version of the Phytophthora Genome Initiative (PGI) database (named the Phytophthora Genome Consortium or PGC), which is expected to be released during 2001.

In addition to cDNA sequencing, coordinated efforts in the genomic sequencing of Phytophthora are emerging. Poster presentations by Stephen Whisson (Scottish Crop Research Institute) and collaborators and by members of the Tyler laboratory described progress in the physical mapping of both the *P. infestans* and *P. sojae* genomes. BAC contigs spanning avirulence gene clusters were described as prime candidates for targeted genomic sequencing.

With the accumulation of sequence data for Phytophthora, the challenge is shifting to functional analyses. Sophien Kamoun (Ohio State University, Wooster) described the use of potato virus X (PVX)-mediated gene expression to carry out functional screens of *P. infestans* genes in plants. The transient nature of the PVX expression system and the use of a recently developed Agrobacterium binary PVX vector (Baulcombe, 1999; Takken et al., 2000) allowed functional analyses of pathogen genes at an unprecedented high throughput rate. By combining EST data mining with PVX-based functional screens, Kamoun and collaborators identified a battery of novel Phytophthora genes that trigger hypersensitive-like necrosis in Nicotiana and tomato.

LESSONS FROM OTHER GENOME PROJECTS

Several talks illustrated the impact of genome-enabled research on the understanding of microbial physiology and evolution. Claire Fraser (The Institute for Genomic Research, Rockville, MD) provided an overview and update of more than 30 microbial genome sequences. In prokaryotes, genomic analyses revealed the tremendous impact of lateral gene transfer on genome evolution. For example, there is evidence for massive gene exchange between bacterial and archaeal thermophiles (Aravind et al., 1998; Nelson et al., 1999). One can expect that with the accumulation of sequence data, the theme of lateral gene transfer among plant-associated microbes and their hosts will become prevalent in the coming years.

Fraser also discussed how the genome sequence of the bacterial pathogen *Neisseria meningitidis*, the cause of human meningitis, was used to identify promising vaccine candidates. This variable pathogen occurs as five pathogenic serogroups, and vaccines are not available against some of the serogroups, such as serogroup B. In a major study published last year, Pizza and coworkers (2000) systematically screened the genome sequence of a virulent strain of *N. meningitidis* serogroup B for surface-associated proteins that are conserved in other *N. meningitidis* strains and in the closely related pathogen *Neisseria gonorrhoeae*. They successfully expressed 350 of the identified ORFs in *Escherichia coli* and used the recombinant proteins to immunize mice. Of these, seven proteins evoked effective antibacterial antibody responses and emerged as prime candidate vaccine antigens. This study clearly illustrates how combining sequence mining with high throughput functional assays can accelerate the identification of genes with desirable characteristics.

MEETING REPORT

Lively discussions centered on the value of partially versus fully sequenced genomes. Fraser and others warned that even with multiple genome coverage, incomplete genome sequences might lack genes that are meaningful to the full understanding of the biology of the examined organism. Nonetheless, incomplete genomes can be useful. For example, mining for surface proteins of *N. meningitidis* was conducted while the sequencing project was still in progress (Pizza et al., 2000). Stanley Malow (University of Illinois, Urbana-Champaign) gave a spirited illustration of how he is using sample sequencing, at 3X to 4X genome coverage, of Salmonella serovars adapted to different animal hosts to gain insight into the genetic basis of host specificity. Differences in host specificity between these serovars are expected to involve multiple loci and do not include characterized pathogenicity gene islands or other known virulence genes. Preliminary analyses of two Salmonella serovars identified ~60 polymorphic insertions and/or deletions that are candidates for functional assays. It will be interesting to see how sample sequencing will help unravel this complex problem and to determine whether similar approaches can be applied to plant-associated microbes that also show a great level of diversity in host specificity.

FUTURE DIRECTIONS

The exponential accumulation of sequence data creates a formidable challenge for the community of plant microbiologists to acquire new skills and expertise in handling and mining large data sets. For example, excellent opportunities exist in mining the sequence data generated through the JGI Microbe Month program for candidate genes relevant to fundamental questions on the association of microbes with plants. It is hoped that several plant microbe research groups will collaborate with bioinformaticists or computer science students to take advantage of the available sequence information.

Some of the sequenced plant microbes interact with model plant species with completed or ongoing genome projects. For example, both *X. c. campestris* and *R. solanacearum* infect the model dicot Arabidopsis, and *S. meliloti* can form symbiotic nodules on the model legume plant *Medicago truncatula*. In the coming years, the integration of microbe and plant genomic data should revolutionize the study of plant–microbe interactions. Of particular interest is the determination of interaction transcriptomes, the sum of transcripts that are produced by the host and the microbe during their association (Birch and Kamoun, 2000). High throughput methods for profiling interaction transcriptomes will allow a comprehensive and coordinated study of gene expression during plant–microbe interactions.

As stated in closing comments by the co-chair of the organizing committee, Peter Johnson, it seems unlikely that the Agricultural Microbes Genome Conference will continue in its current format. Future challenges will be to maintain an identity for the agricultural microbes genome community while coordinating with other microbial genomics gatherings, such as The American Society for Microbiology and The Institute for Genomic Research Conference on Microbial Genomes that was held this year in Monterey, California, from January 29 to 31, 2001. Whatever the future of the Agricultural Microbes Genome Conference will be, the plant microbe community has now fully embraced the era of genomics.

Sophien Kamoun
Department of Plant Pathology

Saskia A. Hogenhout
Department of Entomology
Ohio State University
Ohio Agricultural Research and Development Center

REFERENCES

Andrews, J.H., and Harris, R.F. (2000). The ecology and biogeography of microorganisms on plant surfaces. Annu. Rev. Phytopathol. **38,** 145–180.

Aravind, L., Tatusov, R.L., Wolf, Y.I., Walker, D.R., and Koonin, E.V. (1998). Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. Trends Genet. **14,** 442–444.

Baulcombe, D.C. (1999). Fast forward genetics based on virus-induced gene silencing. Curr. Opin. Plant Biol. **2,** 109–113.

Bertolla, F., Frostesard, A., Brito, B., Nesme, X., and Simonet, P. (1999). During infection of its host, the plant pathogen *Ralstonia solanacearum* naturally develops a state of competence and exchanges genetic material. Mol. Plant-Microbe Interact. **12,** 467–472.

Birch, P.R.J., and Kamoun, S. (2000). Studying interaction transcriptomes: Coordinated analyses of gene expression during plant-microorganism interactions. In New Technologies for Life Sciences: A Trends Guide, R. Wood, ed (London, UK: Elsevier Science), pp. 77–82.

Erwin, D.C., and Ribeiro, O.K. (1996). *Phytophthora* Diseases Worldwide. (St. Paul, MN: American Phytopathological Society Press).

Fenselau, S., and Bonas, U. (1995). Sequence and expression analysis of the *hrpB* pathogenicity operon of *Xanthomonas campestris* pv. *vesicatoria* which encodes eight proteins with similarity to components of the Hrp, Ysc, Spa, and Fli secretion systems. Mol. Plant-Microbe Interact. **8,** 845–854.

Fleischmann, R.D., et al. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269,** 496–512.

Fraser, C.M., et al. (1995). The minimal gene complement of *Mycoplasma genitalium*. Science **270,** 397–403.

Gueneron, M., Timmers, A.C., Boucher, C., and Arlat, M. (2000). Two novel proteins, PopB, which has functional nuclear localization signals, and PopC, which has a large leucine-rich repeat domain, are

MEETING REPORT

secreted through the *hrp*-secretion apparatus of *Ralstonia solanacearum*. Mol. Microbiol. **36,** 261–277.

Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E., and Holden, D.W. (1995). Simultaneous identification of bacterial virulence genes by negative selection. Science **269,** 400–403.

Kamoun, S., Hraber, P., Sobral, B., Nuss, D., and Govers, F. (1999a). Initial assessment of gene diversity for the oomycete pathogen *Phytophthora infestans* based on expressed sequences. Fungal Genet. Biol. **28,** 94–106.

Kamoun, S., Huitema, E., and Vleeshouwers, V.G.A.A. (1999b). Resistance to oomycetes: A general role for the hypersensitive response? Trends Plant Sci. **4,** 196–200.

Kunst, F., et al. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature **390,** 249–256.

Loh, J., Garcia, M., and Stacey, G. (1997). NodV and NodW, a second flavonoid recognition system regulating nod gene expression in *Bradyrhizobium japonicum*. J. Bacteriol. **179,** 3013–3020.

Nelson, K.E., et al. (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature **399,** 323–329.

Pizza, M., et al. (2000). Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science **287,** 1816–1820.

Purcell, A.H., and Saunders, S.R. (1999). Glassy-winged sharpshooters expected to increase plant disease. Calif. Agric. **53,** 26–27.

Qutob, D., Hraber, P.T., Sobral, B.W., and Gijzen, M. (2000). Comparative analysis of expressed sequences in *Phytophthora sojae*. Plant Physiol. **123,** 243–254.

Simpson, A.J., et al. (2000). The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature **406,** 151–157.

Swarup, S., Yang, Y., Kingsley, M.T., and Gabriel, D.W. (1992). An *Xanthomonas citri* pathogenicity gene, *pthA*, pleiotropically encodes gratuitous avirulence on nonhosts. Mol. Plant-Microbe Interact. **5,** 204–213.

Takken, F.L., Luderer, R., Gabriels, S.H., Westerink, N., Lu, R., De Wit, P.J., and Joosten, M.H. (2000). Technical advance: A functional cloning strategy, based on a binary PVX-expression vector, to isolate HR-inducing cDNAs of plant pathogens. Plant J. **24,** 275–283.

Viprey, V., Del Greco, A., Golinowski, W., Broughton, W.J., and Perret, X. (1998). Symbiotic implications of type III protein secretion machinery in *Rhizobium*. Mol. Microbiol. **28,** 1381–1389.

Viprey, V., Perret, X., and Broughton, W.J. (2000). Host-plant invasion by rhizobia. Subcell. Biochem. **33,** 437–456.

Zhu, H., Blackmon, B.P., Sasinowski, M., and Dean, R.A. (1999). Physical map and organization of chromosome 7 in the rice blast fungus, *Magnaporthe grisea*. Genome Res. **9,** 739–750.

Zhu, W. MaGbanua, M.M., and White, F.F. (2000). Identification of two novel hrp-associated genes in the hrp gene cluster of *Xanthomonas oryzae* pv. *oryzae*. J. Bacteriol. **182,** 1844–1853.

Medicago truncatula on the Move!

Plant biology is moving rapidly. With the recent completion of the sequence of the Arabidopsis genome (The Arabidopsis Genome Initiative, 2000), researchers have turned their attention to the genomes of other crop plants (Adam, 2000). Many agronomically important crop plants are legumes, such as soybean, pea, and alfalfa. However, the size and complexity of these genomes makes them unwieldy and has slowed progress on the genetic characterization of these crops. *Medicago truncatula* recently emerged as a model plant for legume genetics and genomics (Cook, 1999). The small diploid genome, autogamous genetics, and ease of transformation make this close relative of alfalfa (*Medicago sativa*) a good model system. The recent development of genetic and genomic tools for *M. truncatula* research has propelled *M. truncatula* into the forefront of legume research as an ideal legume model. Advances in *M. truncatula* genomics were the subject of a recent workshop preceding the 9th Annual Plant and Animal Genome Meeting in San Diego, California, on January 11 and 12.

The international scope of this workshop underscored the interest in *M. truncatula* as a model system for studying plant–microbe interactions, in particular rhizobial and mycorrhizal symbioses. More recently, interest has turned to *M. truncatula* as a system for examining the very rich production of secondary metabolites by legumes and legume-specific disease resistance (Cook et al., 2000; Harrison, 2000). This workshop, which started out three years ago as a small band of National Science Foundation (NSF) genome grant collaborators sponsored by the NSF Plant Genome Project (award number 9,872,664), has grown into an international effort spanning national boundaries and legume species.

This year's participants represented not only the original NSF genome grant collaborators but members of the Institut National de la Recherche Agronomique–Centre National de la Recherche Scientifique (INRA-CNRS), German mycorrhizal and European Union (EU) *M. truncatula* genome projects (funded by the Deutsche Forschungsgemeinshaft and the European Union 5th Framework Program, respectively), the Noble Foundation *M. truncatula* project, and legume scientists working with other species such as pea, soybean, and alfalfa. What emerged from the sharing of data, questions, ideas, and future plans was a coordinated effort using worldwide resources to develop *M. truncatula* as a model plant as easy to use as Arabidopsis with broad applicability to other legume systems.

EST SEQUENCING–THE 100,000 MARK

The most publicly visible advance in the development of *M. truncatula* as a model legume system is the large number of expressed sequence tags (ESTs) deposited in the public domain. The INRA-CNRS-Genoscope, Noble Foundation, and NSF projects all include activities that construct cDNA libraries and sequence ESTs, with the result that *M. truncatula* now contributes approximately as many ESTs to EST databases as soybean and Arabidopsis and likely will soon pass Arabidopsis. Data from three EST projects were presented at the meeting. Pascal Gamas (INRA-CNRS, Toulouse, France) presented data from the INRA-CNRS-Genoscope EST project, Gregory May (Noble Foundation, Ardmore, OK) discussed the Noble Foundation's EST project, and the NSF project was represented by Chris Town (The Institute for Genomic Research, Rockville, MD). Town also presented data on an EST database, the *Medicago truncatula* Gene Index, that integrates all publicly available *M. truncatula* ESTs. Together, this consortium of researchers has sequenced ESTs from 24 cDNA libraries of diverse origins. Not only do the libraries encompass most major plant tissues, but they also incorporate developmental stages and treatments with microbial pathogens, bacterial and fungal symbionts, insect pests, and abiotic stressors. Combining and analyzing the data from several efforts has proven extremely powerful. For example, each EST can be traced to the library from which it was sequenced. By cross-referencing this information with the compilation of homologous ESTs into tentative consensus sequences, the *Medicago truncatula* Gene Index can provide a first indication of expression patterns for the corresponding gene. At the time of the last Gene Index release in December 2000, almost 89,000 *M. truncatula* ESTs were available in public databases, but by the January date of the meeting the number had grown to more than 105,000 and continues to climb. The 88,906 ESTs used to generate the latest release fall into 10,160 tentative consensus sequences comprising 68,844 ESTs and 19,985 singletons, for a total of 30,145 unique sequences (for comparison, Arabidopsis has 34,491 unique sequences in the Gene Index).

DEVELOPMENT OF A UNIGENE SET AND LARGE SCALE EXPRESSION PROFILING

Although the generation of ESTs has value in and of itself, one focus of all the genome projects represented was to produce these sequences for the development of microarrays and markers. Kate VandenBosch (University of Minnesota, Minneapolis/St. Paul), Helge Kuester (Universitaet Bielefeld, Germany), and Gregory May presented progress and future directions for the development of a DNA-chip-microarray resource. VandenBosch presented exploratory array data gathered during the development phase of the chip resource under the NSF Genome Project and discussed progress toward the development of a "universal unigene set." The unigene set will be a defined set of nonredundant cDNA clones that will be made available to the community for use in microarray and other experiments. This will help standardize expression profiling experiments and will allow access to the technology by more researchers. The unigene set will evolve over time, beginning with 6000 genes and growing, depending on future funding, to an expected 20,000 to 30,000 clones covering the breadth of the transcriptome. These genes will be chosen from all of the genes available in the database and sequenced from both ends to confirm identity and obtain complete transcript information where possible. Rod Wing has agreed to add the future EST resource to the *M. truncatula* bacterial artificial chromosome (BAC) libraries that the Clemson University Genomics Institute currently houses and distributes, making the unigene set another publicly available resource. Kuester discussed the EU project's similar goals to use both microarrays and macroarrays, emphasizing the latter as another way to bring the technology to the average user. In addition to transcriptome analysis using microarrays and SAGE (Serial Analysis of Gene Expression), the Noble Foundation has expanded *M. truncatula* expression profiling to include proteomics and metabolic profiling. It has undertaken a large-scale project in *M. truncatula*, using two-dimensional gels and matrix-assisted laser desorption ionization time-of-flight mass spectrometry to identify proteins and electro-spray ionization HPLC mass spectrometry to identify metabolites. Linking these data to the microarray and genome

MEETING REPORT

sequencing data should prove extremely fruitful.

AN EXPONENTIAL LEAP IN THE NUMBER OF T-DNA–TAGGED LINES

Crucial to the development of *M. truncatula* as a model system is the availability of reverse genetic tools. Pascal Ratet (CNRS, Gif sur Yvette, France) and Maria Harrison (Noble Foundation) each discussed the T-DNA–tagged lines being generated by their groups, and Maria Fedorova presented the progress of parallel efforts by a collaboration among Steve Gantt, Debby Samac, and Carroll Vance (University of Minnesota). Ratet's group used a transformation and regeneration protocol (Trinh et al., 1998) and a gene fusion vector to generate 350 T-DNA–tagged lines. Tagged loci have been characterized, and another several thousand lines are planned in the frame of the European *M. truncatula* genome project. Harrison's group developed a vacuum infiltration protocol for *M. truncatula* (Trieu et al., 2000) and combined this with an activation tagging vector (Weigel et al., 2000) to generate 180 lines in a pilot experiment. These lines are currently in the T2 or T3 generation and demonstrate both dominant and recessive phenotypes. Scaling up from the initial experiment, more than 10,000 plants have been infiltrated to date, which should result in between 9000 and 51,000 lines, based on the efficiencies reported by Trieu et al. (2000). Transformation is ongoing at 3000 plants per week. The Minnesota group has conducted a small scale (4000 seedlings) mutagenesis experiment using vacuum infiltration of seedlings and various vectors and plans to perform the same analysis with the promoter trap vectors before embarking on large scale production of tagged lines. When these projects are combined, the number of T-DNA–tagged lines in *M. truncatula* will increase from a few hundred a year ago to a conservative estimate of 125,000 lines within another year.

A COMPREHENSIVE *M. TRUNCATULA* GENETIC MAP

Six years ago, data on the first *M. truncatula* mutant, *TE-7*, was published (Benaben et al., 1995). Since that time, there has been an explosion in the number of mutants isolated and described (Penmetsa and Cook, 1997, 2000; Catoira et al., 2000; Nakata and McConn, 2000; Wais et al., 2000). At the time that *TE-7* was identified, there were few genetic markers and there was no map, genetic or physical. A highlight of this meeting was the presentation of genetic maps by Thierry Huguet (INRA-CNRS, Toulouse, France) and Dongjin Kim (University of California, Davis). The two maps were based on physical differences between different pairs of *M. truncatula* ecotypes, using genotype Jemalong A17 as the recurrent parent. Huguet has identified the location of 240 amplified fragment length polymorphism (AFLP) markers and is currently adding restriction fragment length polymorphisms (RFLPs) and microsatellites to this map. In addition, members of his laboratory, in collaboration with J.M. Prospéri (INRA, Montpellier, France), has generated 195 recombinant inbred lines, 75% of which are at the F7 level. These lines will be available publicly. In a parallel effort, Dongjin Kim reported that the Cook laboratory has mapped 230 codominant polymerase chain reaction markers, 150 of which are linked to clones from a growing BAC genomic library resource (now estimated at 32-fold genome coverage). More than 100 of these markers link ESTs to the map, thus connecting the functional and structural genomics efforts of this group. An ongoing collaboration between the laboratories of Nevin Young (University of Minnesota) and Thierry Huguet to link these two genetic maps is resulting in the identification of microsatellite markers that are polymorphic among all three mapping ecotypes. Integration of a cytogenetic map being constructed by members of Ton Bisseling's laboratory (Wageningen University, the Netherlands) with the genetic map produced by Kim and Cook has helped resolve some ambiguities, allowing 35 BACs to be placed on both maps. These genetic resources will be complemented by a project at the University of California, Davis, where researchers are initiating the development of a complete physical map for *M. truncatula*.

The strength of these extensively linked genetic and physical maps is that after identifying a linked marker, a researcher can go directly to a BAC clone containing genomic DNA from the region of interest. In many cases, these BAC clones have already been linked into a contig, thus eliminating several steps in the walk from phenotype to map position and ultimately to gene identification.

THE STRUCTURE OF THE ARABIDOPSIS GENOME IS AN INSUFFICIENT MODEL FOR LEGUMES

The recent completion of the Arabidopsis genome sequence has provided an enormous boost for the plant field as a whole. However, reports of very low levels of microsynteny between Arabidopsis and *M. truncatula* suggest that the structure of the Arabidopsis genome may not be a good indicator of how legume genomes will look. Using different approaches, Nevin Young reported a global estimate of only 8% microsynteny between Arabidopsis and *M. truncatula* genomes, whereas researchers at the University of California, Davis, estimate ~10% frequency

of microsynteny based on survey sequencing of more than 20 BAC clones in *M. truncatula*. Nevertheless, some genome regions are conserved between these two models, and Jongmin Baek from the Davis group reported detailed sequence analysis of a segmental duplication in *M. truncatula* that exhibits conservation with the Arabidopsis genome.

In addition, an analysis of disease resistance genes by Steve Cannon (University of Minnesota) and Nevin Young in collaboration with Doug Cook has identified resistance genes that have radiated in legumes but that are poorly represented or absent in Arabidopsis. This work suggests that there are some very ancient differences in disease resistance genes and that Arabidopsis may be a poor model for studying disease resistance in the crop legumes.

M. TRUNCATULA AS A NODAL SPECIES FOR COMPARATIVE LEGUME GENOMICS

Influenced by the fact that Arabidopsis may not be a good model for the structure of legume genomes, a large focus of this meeting was on possible synteny between *M. truncatula* and other legumes, with the goal of ultimately developing a composite legume genomic map. Attention centered on three agronomically important legumes: alfalfa, pea, and soybean. Gyorgy Kiss (Hungarian Academy of Sciences, Szeged, Hungary) and Dongjin Kim have placed more than 60 markers on a comparative map linking *M. truncatula* with its close relative alfalfa, demonstrating a high level of macrosynteny between these two species.

A survey of microsynteny between soybean and *M. truncatula* starting with ~20 Mbp of the soybean genome (50 BAC contigs) found very high levels of microsynteny between the two genomes. Nevin Young reported that 54% of the soybean BACs exhibited some microsynteny with *M. truncatula*, of which most (>70%) exhibited extensive conservation. These observations suggest that information derived about the *M. truncatula* genome has enormous implications for soybean.

Evidence for such high levels of microsynteny between *M. truncatula* and other legumes has prompted the development of cross-genome markers that would link the genetic maps of several important legume species. Dongjin Kim reported on progress in developing cross-genome markers for *M. truncatula*, pea, mung bean, soybean, and *Lotus japonicus*. The idea behind such a composite map linking these genomes is that researchers studying a gene in one legume species could easily move to another species to map and clone it and then return to the original species for analysis. At 450 Mbp, the *M. truncatula* genome is only a few times larger than that of Arabidopsis, but it is 10 times smaller than that of pea. Unlike alfalfa and soybean, *M. truncatula* is a diploid and thus amenable to genetic analysis. The ease of genetic analysis in *M. truncatula* and its more manageable genome size makes it the species of choice for mapping and cloning. In concept, *M. truncatula* would be the nodal species that would link the genetic maps of the larger legume genomes together. In addition, these cross-genome markers would link these maps to that of *L. japonicus*, another legume species that has been the focus of a large molecular genetic effort.

IDENTIFICATION OF GENE-RICH REGIONS

One of the striking features of legume biology is the variation in genome size among closely related species. The pea genome is almost 10 times the size of the *M. truncatula* genome, even though their haploid chromosome number differs by only one. A comparison of the pea, alfalfa, and *M. truncatula* genomes by Noel Ellis (John Innes Centre, Norwich, UK), Gyorgy Kiss, and members of the Cook laboratory demonstrates that the difference in size between the pea and *Medicago* genomes is not due to genome duplication. Rather, the large size of the pea genome seems to be due to the recent accumulation of retrotransposons and noncoding regions of DNA (Noel Ellis).

These observations raised the issue of gene-rich regions. Identification of such regions could increase the productivity of a cloning or massive sequencing effort. In the process of creating a cytogenetic map, members of the Bisseling laboratory have determined that the genome of *M. truncatula* has a remarkably efficient organization. One estimate suggests that ~80% of the genome is organized into pericentromeric heterochromatin, whereas the chromosome arms are composed primarily of euchromatic DNA. Thus, the gene-rich region of the genome may be condensed into as little as 100 Mbp of DNA.

Sequencing the *M. truncatula* genome may not be far off. Jean Denarié (INRA, Toulouse, France) reported on recent international workshops held in Europe and the United States to identify priorities in the use of *M. truncatula* to facilitate the breeding of crop legumes (Cook and Denarié, 2000) and the possibility of a massive *M. truncatula* sequencing effort of international proportions. A follow-up meeting has been planned for the spring to discuss strategic plans.

BIOINFORMATICS RESOURCES FOR *M. TRUNCATULA*

To keep pace with the rapid increase in *M. truncatula* genomics, several

MEETING REPORT

bioinformatics groups have established searchable databases that assemble *M. truncatula* EST and/or BAC sequences, report homologies, and are freely accessible through the World Wide Web. The National Center for Genome Resources (Santa Fe, NM) in collaboration with the Noble Foundation has established the Medicago genome initiative site at http://www.ncgr.org/research/mgi/ (Bell et al., 2001). Callum Bell (National Center for Genome Resources) pointed out that, as in other informatics fields, Medicago informatics researchers are grappling with how to reduce redundant effort, break down barriers that separate data types, and cross species boundaries. Bringing the informatics leaders from several projects together at this meeting was seen as a step toward solving these problems. The EU and INRA data are available at http://sequence.toulouse.inra.fr/Mtruncatula.html, with a searchable index under development. The NSF project has fostered the development of two databases: The Institute for Genomic Research *Medicago truncatula* Gene Index, which is accessible at http://www.tigr.org/tdb/mtgi/, and a species database site (MtDB) at http://chrysie.tamu.edu/medicago/. Ernie Retzel (University of Minnesota) presented a new master link to Medicago-related sites under development at http://www.medicago.org/. This site also will contain comparative information for the soybean project housed at the University of Minnesota as well as protein family assignments. These sites are being updated regularly and are proving to be an invaluable tool for *M. truncatula* genetics and genomics research.

UPCOMING EVENTS

Perhaps the most important outcome of the meeting was the heightened interactions between the research groups, large and small, working on *M. truncatula*. The international nature of the effort should be evident this summer at the 4th Workshop on *M. truncatula*, a satellite meeting of the International Molecular Plant–Microbe Interactions conference (http://www.plantpath.wisc.edu/mpmi/satellite.html). In Europe, a European Molecular Biology Organization practical course in November 2001, at Gif sur Yvette, France will introduce the system to new researchers (http://www.isv.cnrs-gif.fr).

CONCLUSIONS

This conference highlighted the recent explosion in the development of genetic and genomic tools for *M. truncatula*. The reports of the sequencing of more than 100,000 ESTs, extensive genetic maps linked to BAC genomic clones, progress in activation tagging, and the initiation of large-scale T-DNA mutagenesis demonstrate that the genetic and genomic tools are finally in place to use *M. truncatula* as a model legume. This strategic focus on a tractable legume and the development of a composite legume genetic map will enable researchers to move easily between *M. truncatula*'s simple genome and those of other legumes. With the rapid advances described above and the possibility of initiating a genome sequencing effort, *M. truncatula* will not be just on the move, it will be flying!

ACKNOWLEDGMENTS

We thank Doug Cook for organizing the meeting and all of the conference participants for their contributions in the lecture and discussion sessions and for sharing their presentation materials with us. A list of meeting attendees is available on the World Wide Web at http://www.medicago.org/documents/Meeting_documents/SanDiego_part.html.

Julia Frugoli
Department of Biological Sciences/Biochemistry
Clemson University
132 Long Hall
Clemson, SC 29634

Jeanne Harris
Department of Botany and Agricultural Biochemistry
Marsh Life Science Building
University of Vermont
Burlington, VT 05405-0086

REFERENCES

Adam, D. (2000). Now for the hard ones. Nature **408,** 792–793.

Bell, C.J., Dixon, R.A., Farmer, A.D., Flores, R., Inman, J., Gonzales, R.A., Harrison, M.J., Paiva, N.L., Scott, A.D., Weller, J.W., and May, G.D. (2001). The Medicago Genome Initiative: A model legume database. Nucleic Acids Res. **29,** 114–117.

Benaben, V., Duc, G., Lefebvre, V., and Huguet, T. (1995). TE7, an inefficient symbiotic mutant of *Medicago truncatula* Gaertn. cv. Jemalong. Plant Physiol. **107,** 53–62.

Catoira, R., Galera, C., de Billy, F., Journet, E.P., Maillet, F., Penmetsa, V., Rosenberg, C., Gough, C., Cook, D., and Denarié, J. (2000). Identification of four genes of *Medicago truncatula* controlling steps in Nod factor transduction. Plant Cell **12,** 1647–1665.

Cook, D. (1999). *Medicago truncatula*: A model in the making! Curr. Opin. Plant Biol. **2,** 301–304.

Cook, D., and Denarié, J. (2000). Progress in the genomics of *Medicago truncatula* and the promise of its application to grain legume crops. Grain Legumes **28,** 12–13.

Cook, D., Kim, D.J., Zhu, H.Y., and Uribe, P. (2000). Plant–pathogen interactions in *Medicago truncatula*. Grain Legumes **28,** 20.

Harrison, M.J. (2000). Molecular genetics of model legumes. Trends Plant Sci. **5,** 414–415.

MEETING REPORT

Nakata, P.A., and McConn, M.M. (2000). Isolation of *Medicago truncatula* mutants defective in calcium oxalate crystal formation. Plant Physiol. **124,** 1097–1104.

Penmetsa, R.V., and Cook, D. (1997). A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. Science **275,** 527–530.

Penmetsa, R.V., and Cook, D.R. (2000). Production and characterization of diverse developmental mutants of *Medicago truncatula*. Plant Physiol. **123,** 1387–1397.

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

Trieu, A.T., Burleigh, S.H., Kardailsky, I.V., Maldonado-Mendoza, I.E., Versaw, W.K., Blaylock, L.A., Shin, H., Chiou, T.J., Katagi, H., Dewbre, G.R., Weigel, D., and Harrison, M.J. (2000). Transformation of *Medicago truncatula* via infiltration of seedlings or flowering plants with *Agrobacterium*. Plant J. **22,** 531–541.

Trinh, T.H., Ratet, P., Kondorosi, E., Durand, P., Kamate, K., Bauer, P. and Kondorosi A. (1998). Rapid and efficient transformation of diploid *Medicago truncatula* and *Medicago sativa* ssp. *falcata* lines improved in somatic embryogenesis. Plant Cell Rep. **17,** 345–355.

Wais, R.J., Galera, C., Catoira, R., Penmetsa, R.V., Cook, D., Gough, C., Denarié, J., and Long, S.R. (2000). Early symbiosis mutants of *Medicago truncatula* show a block in Ca^{++} spiking. Proc. Natl. Acad. Sci. USA **97,** 13407–13412.

Weigel, D., et al. (2000). Activation tagging in Arabidopsis. Plant Physiol. **122,** 1003–1014.