and their starting positions were adjusted using RBSfinder. The ORFs were then subjected to an in-house semi-automatic genome annotation system, which is based on pairwise/HMM searches against public databases. Of the putative protein-coding genes, 1,621 (33.1%) could be functionally assigned by COG, and 2,193 (45.3%) showed significant matches to characterized protein sequences in the databases. 1,025 ORFs (21.6%) have no significant similarities to proteins in the databases.

The genome sequence contains at least 15 sigma factors. Ten of them belong to the ECF family of sigma factors. 108 genes possibly involved in the sporulation process were identified based on their amino-acid sequence identities of more than 30% to known sporulation genes of other bacilli and clostridia. Other interesting groups of genes present in the E681 genome are those for production of polyketides, lantibiotics or non-ribosomal peptides, which may function as siderophores or antimicrobials. E681 appears to produce at least five kinds, since polyketide synthase genes are found in one supercontig, lantibiotic genes in one supercontig and non-ribosomal peptide synthetase genes are in four supercontigs.

Approaches of functional genomics including systematic mutagenesis, DNA microarray and proteome analysis are being applied to E681 to understand the regulation of developmental processes such as sporulation, metabolic processes and the nature of the 'plant-probiotic' property. Random mutagenesis of the E681 strain using mini-Tn*10* transposon was carried out successfully and more than 15,000 mutants have been generated. We obtained 180 sporulation-deficient mutants and 23 of them were analyzed; the flanking region of mini-Tn*10* transposition of each mutant was recovered by plasmid rescue and the nucleotide sequence was determined. The mutations were found on the homologues of *sigE*, *spoIIIAH*, *asnO*, and some unknown protein genes, respectively. We analyzed 86 mini-Tn*10* mutants lost antifungal activities against *R. solani* and/or *F. oxysporum*, and mutations were found on the homologues of a polykitide synthase gene and other 18 different genes. We also obtained 17 mutants showing changed auxin activities and analyzed the mutation sites. Two mutant strains carrying mutations on a homologue of a transcriptional regulator gene and an unknown gene, respectively, were analyzed to produce 15-times and 17-times higher auxin than mother strain, respectively.Also, for the monitoring of transcriptome profiles of E681 we selected candidate genes which seem to be involved in developmental processes such as sporulation, global regulatory system, synthesis of antimicrobial compounds, and built a 0.5K-scale experimental oligonucleotide microarray chip. Using this chip we are analyzing the transcriptome of E681 cells grown together with a fungal plant pathogen. 2-D gel electrophoresis and MALDI-TOF analysis are being set up to identify proteins produced during interaction with plants. We also found that *P. polymyxa* E681 produces antifungal antibiotics, fusaricidins A and B, by HPLC and LC/MS analysis.

In another effort, we isolated forty *Paenibacillus* spp. closely related to the *P. polymyxa* E681 strain from rhizosphere. The genome sizes of forty *Paenibacillus* isolates were measured with pulsed-field gel electrophoresis using the genomic DNA digested with *Not*I, *Pme*I and *Sfi*I. The estimated genome sizes were ca. 5.2-5.5 Mb. Five distinct groups were identified based on large-restriction-fragment (LRF) polymorphism of *Not*I- and *Pme*I-digested isolates. About half of the isolates were in a single group. The presence of absence of *nifH*, plasmid profiles, and results of box PCR further supported this grouping.

SI-4

Systems biology initiatives in the rice blast fungus, *Magnaporthe grisea*

Yong-Hwan Lee

School of Agricultural Biotechnology, Seoul National University, Seoul 151-921, Korea.

Rice blast, caused by *Magnaporthe grisea*, has been considered as a model plant disease to study plant-fungi interactions. This is due to not only economic significance of this disease worldwide but genetic and molecular tractability of the fungal pathogen. These features include genetic crossing with two different mating types, extensive genetic maps, developments of transformation and gene knock-out technologies. Extensive research has been conducted to understand infection mechanisms by the pathogen and defense mechanisms of the host at cellular and molecular biology levels during the last decade. Through an elegant series of research, we know environmental cues and related signaling systems involved in infection of host plant by the fungus. More than a dozen of genes are identified as pathogenicity determinants through insertional mutagenesis using REMI (restriction enzyme mediated integration) or reverse genetic strategy. However, the precise mechanisms to complete the disease cycle remain to be understood. Recent advances on genomics research are making much progress to approach the ultimate understanding of pathogenesis at biochemical and molecular biological levels. In 2002, both whole genome drafts of rice and *M. grisea* were obtained and all information is available in public. Currently much effort is being focused on accurate annotation of genes in both organisms.

Agrobacterium tumefaciens-mediated transformation (ATMT) has long been used to transfer genes to a wide variety of plants and has also been used extensively as a tool for insertional mutagenesis in *Arabidopsis thaliana*. More recently, several fungi have been transformed using *A. tumefaciens*. For

insertional mutagenesis, this technique offers huge potential as an alternative tool to REMI. One of the principal advantages of ATMT over conventional transformation techniques is the versatility in choosing which starting material to transform. Applicable to several fungi, *A. tumefaciens* can transform protoplasts, hyphae, spores, and blocks of mushroom mycelial tissue. Furthermore, ATMT generates a high percentage of transformants with a single insert of T-DNA in the fungal genomes, which will facilitate the subsequent isolation of tagged genes especially from those fungi lacking a sexual stage. Recently we developed ATMT technology in *M. grisea*.

To dissect the function of potential fungal pathogenicity genes in the level of genome-wide in the rice blast fungus, we initiated this project that includes 1) generation of mutants using ATMT 2) development of high throughput screening technology for mutants 3) high throughput technology for DNA extraction, and finally 4) rescuing flanking sequences from interested mutants. Moreover, we are establishing infrastructure for bioinformatics tools for systems biology.

Fungal strain of *M. grisea*

We measured genetic make-up of *M. grisea* KJ201, a Korean field strain, using different transposons, avirulence genes, growth rate, and mating ability. Strain KJ201 is quite different from strain 70-15 that whole genome was sequenced. Strain KJ201 contained much less copy number of transposon MAGGY and telomere linked helicase. Avirulence genes (*PWL2* and *Avr-Pita*) are present in both strains but the copy number was different. Strain KJ201 is female sterile whereas 70-15 is hermaphroditic in their fertility.

Development of high throughput screening technology (HTS)

To characterize phenotypes of many transformants (mutants) at the same time, development of high throughput screening (HTS) system is required. The concept and technology of HTS have been developed and widely used in pharmaceutical area to evaluate new potential candidates for new drugs but not well developed in plant pathology field. One of bottle necks is lack of specific sites we can apply to. However, there exists HTS system in certain aspects in pesticide developing companies.

All single-spored mutants were grown in 24-well tissue culture plate to avoid labeling each mutant number. Further we grow all mutants in liquid medium in 24-well plate to harvest mycelia. By changing media compositions, we can use this technology to measure conidiation ability. In this program, we measured growth rate, pigmentation, conidial morphology by direct examination on 24-well format. Conidiation ability, conidial germination and appressorium formation of all transformants were measured under microscope.

High throughput pathogenicity assay

Traditionally plants are grown in soil in pots to test fungal pathogenicity. However, it is not always easy to provide the same environmental conditions every time. As we all aware, disease development is the outcome of combination of host plant, pathogen, and environment. We developed the protocol for growing rice plants on tissue culture medium in the laboratory. Further we removed rice seed coats. Removing seed coats give several advantages including efficient surface sterilization, more homogeneous germination and speeding up plant growth. Further we can control the concentration or composition of tissue culture medium and ingredients as well. It has been understood that most rice cultivars become more susceptible when more nitrogen is supplied in their culture medium (soil). Conidia harvested from V-8 juice medium on 24-well format were directly sprayed into rice seedlings grown on test tubes. Disease rating was performed 6 days after inoculation.

Mutants selected for further study

We obtained more than two thousands of transformants showing more than one defect from primary HTS procedure. These selected mutants were further tested for their pathogenicity defects on rice seedlings. Once their defects on pathogenicity is verified on rice seedlings, these mutants were grown in liquid medium and genomic DNAs were isolated. By Southern blot analysis, each mutant was verified for T-DNA insertion on the chromosome and its copy number. For large scale DNA extraction from transformants, we also developed high throughput technology. Fungal transformants were grown on liquid medium on 24-well plate until they depleted the medium. Then the plates were freeze-dried and kept until use. Genomic DNA extraction was performed directly from 24-well plate.

Identification of flanking sequences of T-DNA insertion

Mutants, pathogenicity defective ones, verified as single integration of T-DNA were subjected to TAIL-PCR to rescue flanking sequences. Since genome draft of *M. grisea* is available, we could identify full ORF sequences once we get a part of sequences on insertion sites. Thus far, we identified more than 250 loci from the genome sequence. Further we also investigated the integration pattern of T-DNA into fungal genome.

Development of infrastructure for systems biology

In addition to genomics works, we constructed relational databases to manage huge genomics data. Presently, a web-based portal system (CROSS-FPP) is established for efficient data acquisition and analyses such as BLAST and INTERPRO searches, gene ontology (GO) classification, expression profiling, phylogenetic and comparative analyses, and tools for further functional genomics research. This portal system also include all information management for ATMT transformants.