Workshop IX

Fungal Genomics
Chair: Peter Philippsen

IXo-1

SYNTENY PATTERNS: A MESSAGE FROM THE PAST FOR TODAY'S INTERPRETATION OF FUNGAL GENOMES.

Sophie Brachat, Fred Dietrich*+, Sylvia Voegeli*, Riccarda Rischatsch*,*
Iza Kaminski and Peter Philippsen**

**Applied Microbiology, Biozentrum, University of Basel, Switzerland,*
+Department of Molecular Genetics and Microbiology, Duke University, Durham, NC, USA, Phone: ++41 61 267 14 80, Fax ++41 61 267 14 81, peter.philippsen@unibas.ch

We have sequenced the genome of the filamentous ascomycete Ashbya gossypii and produced a complete annotation of the 4718 protein coding genes. (GenBank accession numbers AE016814-AE016821). The systematic gene nomenclature follows that used for Saccharomyces cerevisiae. This facilitated the alignment of homologous regions between both fungal genomes to determine the degree of synteny, and it allowed us to identify formerly consecutive gene orders which were interrupted by genome rearrangements. The complete synteny map reveals that always two S. cerevisiae gene regions show homology and relaxed synteny to A. gossypii genes and that this so called double synteny pattern covered close to 98 % of the A.gossypii genome. This result indicates that both organisms have evolved from the same ancestor, and it provides definite evidence for a genome duplication in the evolution of S. cerevisiae, e.g. by fusion of two related species.

We counted 328 double breaks of synteny in the synteny map of all seven
A. gossypii chromosomes. These are positions in the map where gene orders of both aligning S. cerevisiae regions are interrupted indicating a translocation or inversion event in the A. gossypii lineage or in the precursor of S. cerevisiae prior to its genome duplication. We also counted 168 single breaks of synteny where the gene order in one aligning S. cerevisiae region was interrupted. These sites indicate a translocation or inversion event in the S. cerevisiae lineage after genome duplication. The synteny analysis also revealed that long gene regions in A. gossypii carrying up to 180 genes did not rearrange during 100–200 million years of evolution.

A gene by gene comparison in the synteny map finally shows, based on inferred ancient gene orders, which of the duplicated genes lost one copy and which did not. Combined with the identification of genome rearrangement break points this enables us for example to draw conclusion about non-rearranged or potentially rearranged promoter regions for orthologous genes. Such alteration in promoter sequences may be one of the determinants of fungal diversity.

IXo-2

FUNCTIONAL GENOMIC STUDIES OF ASPERGILLUS FUMIGATUS PATHOGENESIS USING DNA MICROARRAYS

Heenam Kim[1], Christine E. Shamblin[1], Dan Chen[1], Yan Yu[1], Saul H. Sarria[1], Jennifer R. Wortman[1], Jiaqi Huang[1], Resham Kulkarni[1], Gregory S. May[2], David Denning[3], Michael Anderson[3], Aimee K. Zaas[4], and William C. Nierman[1]

[1]The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20854, USA
[2]The University of Texas M. D. Anderson Cancer Center, Houston, TX USA
[3]University of Manchester, Oxford Road, Manchester M13 9PT, UK
[4]Division of Infectious Diseases and International Health, Duke University Medical Center, Durham, NC 27705, USA

Aspergillus fumigatus is an opportunistic pathogen that can cause allergic responses to sensitive individuals and, more seriously, invasive diseases in immuno-compromised individuals. To initiate genomic studies on this pathogen, we have constructed a whole genome DNA microarray. We used a strategy which does not require cDNA libraries or heavily rely on accurate genome annotation, because expressed sequence tag (EST) and cDNA resources of Aspergillus are still very limited, and therefore accurate prediction of the gene structures is difficult. We selected a 700-bp region immediately upstream of the predicted stop codon from each gene. In the case the gene is smaller than 700 bp, we took the entire gene. Then, we included 150 bp of sequence downstream of the gene or as much as there is in the intergenic region when it is shorter than 150 bp. These comprise the target sequences which provide a minimum of 850 bp for each gene. We conducted automated selection of PCR primer pairs by feeding the target sequences to Primer 3.0, with optimized design parameters. Using this approach we were able to design primers for 9,516 of the 9,544 predicted genes (99.7%). We amplified these target gene regions from genomic DNA. The resulting PCR products were purified and spotted in triplicates at high density on aminosilane-coated microscope slides. We are conducting experiments using this microarray to provide insights into growth, virulence mechanisms, and antifungal drug resistance in A. fumigatus.

Presenting author Tel: (301)838-5846
Fax: (301)838-0208
e-mail: hkim@tigr.org

IXo-3

PROTEOME ANALYSIS IN *TRICHODERMA REESEI*

Tiina M. Pakula, Tero Riipi[1], Heini Koivistoinen, Karin Lanthaler[2], Markku Saloheimo, Leena Valmu[1], Nisse Kalkkinen[1], Geoff Robson[2] and Merja Penttilä

VTT Biotechnology, P.O.Box 1500, 02044 VTT, Finland, Tel. +358-9-456 78628, Fax. +358-9-455 2103, email: Markku.Saloheimo@vtt.fi

[1] Institute of Biotechnology, University of Helsinki, Finland

[2]University of Manchester, School of Biological Sciences, 1800 Stopford Building, Oxford Road, Manchester M13 9 PT, UK.

We have set up methods for intracellular and extracellular proteome analysis in Trichoderma. These include analysis of steady state total protein pattern using 2D gel electrophoresis as well as detection of actively synthesised proteins using metabolic labelling. In addition, phosphoprotein staining and Western analysis with phosphoprotein antibodies have been used to detect differences in the cellular levels of phosphorylated proteins.

Proteome analysis has been applied to study the cellular responses activated in Trichoderma reesei upon production of a heterologous protein, tPA (tissue plasminogen activator). Protein samples from chemostat cultures of the tPA-producing transformant and the parental strain Rut-C30 were subjected to 2D gel electrophoresis and protein spots with altered intensity were identified using LC-MS/MS analysis and comparison of the obtained peptide masses and sequence tags with public protein sequence data bases. The analysis revealed up-regulation of a number of proteins involved in protein glycosylation and folding, and members of heat shock protein families. Many of these proteins have not been previously reported from T. reesei. The responses detected in the strain producing the heterologous protein were also compared to those observed in cultures treated with chemical agents to inhibit protein folding and transport, dithiothreitol (DTT) and Brefeldin A, respectively.

IXo-4

A PROTEOMIC COMPARISON OF DIFFERENTIALLY EXPRESSED SECRETED PROTEINS OF ASPERGILLUS FLAVUS, USING TWO-DIMENSIONAL GEL ELECTROPHORESIS (2DE), MALDI-TOF MASS SPECTROMETRY AND TANDEM MASS SPECTROMETRY

Medina, Martha L.[*], Kiernan, U.A.[‡], Haynes, P.[†] and Francisco, W.A.[*]

[*]Arizona State University, Department of Chemistry and Biochemistry, Tempe, AZ 85287-1604, USA; Phone: (480) 965-2403; Fax: (480) 965-2747; e-mail: Martha.Medina@asu.edu, Wilson.Francisco@asu.edu

[‡]Intrinsic Bioprobes Inc. 625 S. Smith Rd. Suite 22, Tempe, AZ 85281, USA

[†]University of Arizona, Biochemistry and Molecular Biophysics, Tucson, AZ 85721-0088, USA

Session: Fungal Genomics; if selected for oral presentation—PowerPoint

Proteomic studies in filamentous fungi are severely lacking. This is particularly apparent for secreted proteins expressed via dispensable metabolic pathways. New techniques in proteomics are making possible the identification of proteins from mass spectrometry data, even those from organisms whose genomes have not been sequenced. Peptide fingerprints and sequences obtained by MALDI-TOF MS and tandem mass spectrometry, respectively, are used to identify proteins using a database of protein sequences derived from genome projects.

Aspergillus flavus is capable of degrading the flavonoid rutin (quercetin 3-glycoside), as the only source of carbon via an extracellular enzyme system. In this continuing study, a proteomic analysis was used to identify secreted proteins from A. flavus when grown on rutin. The growth media glucose and potato dextrose were used as comparisons to identify differentially expressed proteins. The secreted proteins were analyzed by 2DE, MALDI-TOF MS and tandem mass spectrometry. The proteins identified to date will be presented along with a more in-depth characterization of some of the proteins identified, including quercetinase, the enzyme responsible for breaking down quercetin, and several enzymes that have shown fibrinolytic properties. This study has been done in order to gain a more complete understanding of the metabolism of rutin, as well as to expand the knowledge of proteins and enzymes secreted by Aspergillus flavus under varying conditions of growth.

Discovery of secreted proteins and peptides during fungal-plant interactions
by Lene Lange and Kirk Schnorr, Microbial Discovery, Novozymes

Lene Lange, professor, dr.scient., Science Director, Microbial Discovery, Novozymes A/S
Smørmosevej 25, DK2880 Bagsværd, Denmark, 45 4442 3717, 45 4442 1408
lla@novozymes.com

A new discovery method has been developed with the objective of finding secreted enzymes with unknown/undefined activity. It builds on direct selection in live cells (*E.coli*). It is named: Transposon assisted signal trapping (TAST).The method was designed to discover secreted proteins with special emphasis on discovery of enzymes with unknown function. It has been demonstrated that this method has the additional advantage of being extremely efficient in discovery of the proteins and peptides secreted as signals during eukaryote interactions as e.g. taking place between plants and fungi.

TAST has the following procedural steps:
1. Construction of gene library and isolation of plasmids from a pool of >10.000 clones to ensure a certain representivity of the library
2. Transposition reaction with SigA2 (contains a chloramphenicol resistance gene and a signal-less β-lactamase gene)
3. Electroporation into *E. coli* and selection on CAM/Kan (to calculate overall transposition efficiency) and CAM/Kan/Amp (signal trapped clones)
4. Plasmid isolation from ca. 400 CAM/Kan/AmpR clones (trappants) and sequence with primers specific for the transposon *OBS The clue is that the b-lactamase gene with no signal peptide was cloned into the muA minitransposon* such that in frame fusions with a functional signal peptide result in ampicillin resistance of the E.coli colony!*
5. Automatic contig assembly of the sequences and auto annotation
6. Manual analysis of ORF

Advantages of TAST for studying fungal-plant interaction:
cDNA libraries can be constructed, of infected plant materials in the stage of interest
Signal trapping of such composite cDNA libraries gives a short cut to discovery of the proteins and peptides expressed during host-pathogen interaction.
Based on studies of libraries of the plant and fungus separately and through bioinformatics the genes specifically expressed during plant-fungus interaction can be identified
The big surprise has been that *E.coli* can recognize signals of all types of fungi, as well as of Oomycete genes and of plant genes.

GENOME SEQUENCING OF *ASPERGILLUS ORYZAE*

Masayuki Machida[1], Motoaki Sano[1], Kiyoshi Asai[1], Toshitaka Kumagai[1], Taishin Kin[1], Hideki Nagasaki[1], Takashi Komori[2], Toshihiro Tanaka[3], Rie Igarashi[3], Toshihiko Sawano[3], Hisashi Kikuchi[3], Toshihide Arima[4], Osamu Akita[4], Kenichi Kusumoto[5], Yutaka Kashiwagi[5], Keietsu Abe[6], Katsuya Gomi[6], Michio Takeuchi[7], Tetsuo Kobayashi[8], Hiroyuki Horiuchi[6], Katsuhiko Kitamoto[9], Jiujiang Yu[9], Deepak Bhatnagar[9], Thomas E. Cleveland[10], A. oryzae genome sequencing consortium of Japan[11] [[1]Natl. Inst. of Advanced Ind. Sci. and Technol. (AIST) (Higashi 1-1-1, Tsukuba, Ibaraki 305-8566 JAPAN, p:+81-29-861-6164, f:+81-29-861-6174 / Aomi Frontier Building 17F, 2-43 Aomi, Koto-ku, Tokyo 135-0064 JAPAN, p:+81-3-3599-8089, f:+81-3-3599-8085), [2]INTEC Web and Genome Informatics (1-3-3 Shinsuna Koto-ku Tokyo 136-8637, JAPAN, p:+81-3-5665-5141, f:+81-3-5665-5004), [3]Natl. Inst. of Technol. and Eval. (Nishihara 2-49-10, Shibuya-ku Tokyo 151-0066, JAPAN, p:+81-3-3481-8423, f:+81-3-3481-8424), [4]Natl. Res. Inst. of Brewing (3-7-1 Kagamiyama, Higashihiroshima, Hiroshima 739-0046, JAPAN, p:+81-824-20-0823, f:+81-824-20-0808), [5]Natl. Food Res. Inst. (2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, JAPAN, p:+81-29-838-8077, f:+81-29-838-7996), [6]Tohoku U. (1-1 Tsutsumidori-Amamiyamachi Aoba-ku, Sendai, 981-8555 JAPAN, p:+81-22-717-8901, f:+81-22-717-8902), [7]Tokyo U. of Agric. Technol. (Saiwaityou 3-5-8, Fuchu, Tokyo 183-0054, JAPAN, p:+81-423-67-5707, f:+81-423-67-5715), [8]Nagoya U. (Furo-cho, Chikusa-ku, Nagoya 464-8601, JAPAN, p:+81-52-789-4085, f:+81-52-789-4087), [9]The U. of Tokyo (Yayoi 1-1-1, Bunkyo-ku, Tokyo 113_8657, JAPAN, p:+81-3-5841-5161, f:+81-3-5841-8033), [10]South Regional Res. Center, (1100 Robert E. Lee Blvd., New Orleans, LA, 70124, USA, p:+1-504-286-4405, f:+1-504-286-4419), [11]Brewing Soc. of Japan (representative) (2-6-30 Takinogawa, Kita-ku, Tokyo 114-0023, JAPAN, p:+81-3-3910-3853, f:+81-3-3910-3748)]

The genome sequencing project for *Aspergillus oryzae* launched in August 2001 by the whole genome shotgun approach. The contigs were assembled from 6X depth of coverage of total sequence and were connected by end sequencing of 5,000 cosmid clones, yielding approximately 30 scaffolds. Most of the scaffolds have already been mapped on the chromosomes and the relative position of the scaffolds on the chromosomes have been almost completed. There are approximately 20 physical gaps remaining right now, most of which will be closed by BAC libraries. The total genome size of *A. oryzae* was estimated to be 36.8 Mb, which was slightly bigger than that estimated by PFGE. We have already obtained 15 independent telomere sequences, indicating most of the chromosomal ends have been identified. However, we found that the longest band on PFGE might include two chromosomes (I and I'). Instead, the shortest band originally predicted to have two chromosomes (VII and VIII) might have only one. The computational prediction of genes by GeneDecoder (Asai et al., http://www.cbrc.jp/) in combination with alignment with *A. oryzae* ESTs suggested existence of more than 11,000 genes in *A. oryzae* genome, among which approximately 40% of genes were predicted to have introns. As was suggested from the *A. oryzae* genome size bigger than that of *Aspergillus nidulans*, there were many redundant genes including energy supplying enzymes, sugar transporters, hydrolytic enzymes found in the *A. oryzae* genome although the genes encoding RNA polymerase remained single. The analyses of genes for transcription factors, metabolic pathways and so on, which are important especially for industrial application of *A. oryzae*, are in progress.

IXo-7

THE WHOLE GENOME SEQUECE OF THE WHEAT AND BARLEY PATHOGEN, FUSARIUM GRAMINEARUM.

H. Corby Kistler[1], Bruce Birren[2], Li-Jun Ma[2], Sarah Calvo[2], James Galagan[2], Liane R. Gale[1], Kerry O'Donnell[3], Frances Trail[4], Todd Ward[3], Jin-Rong Xu[5] and the Gibberella zeae International Genomics Initiative. [1]USDA ARS Cereal Disease Laboratory and University of Minnesota, St. Paul, MN, USA. Phone: 651-638-1992, FAX 612-625-5054, email, hckist@umn.edu; [2] Broad Institute, Center for Genome Research, MIT, Cambridge, MA, USA; [3]Microbial Genomics Research Unit, USDA ARS, Peoria, IL, USA; [4]Michigan State University, East Lansing, MI, USA; [5]Purdue University, West Lafayette, IN, USA.

We have generated a draft sequence assembly of the F. graminearum genome that is available on the web for download and query. The sequence is of high quality with the entire 36 Mb assembly consisting of just 511 contigs (> 2 kb) contained within 28 supercontigs. The second genome release (October 2003) contains automated annotation, preliminary genome analysis and integration with the genetic map. Using organism-specific parameters for gene prediction, 11,640 protein-coding genes have been identified, representing over 1,500 more genes than predicted by the same method for the non-pathogenic filamentous fungi, Neurospora crassa and Aspergillus nidulans. A genetic map, currently consisting of 66 SNPs, 29 RFLPs, and 27 microsatellites, has been constructed that anchors 99.5% of the sequence assembly. Details of the automated annotation, efforts toward manual annotation and coordination of functional analysis of the genome will be discussed. The F. graminearum sequencing project was funded by the National Research Initiative, through the USDA/NSF Microbial Genome Sequencing Program.