Microbial and Plant Genomics Institute Annual Retreat

Saturday
April 19, 2008

Cowles Auditorium
Hubert H. Humphrey Center

MPGI
MICROBIAL & PLANT GENOMICS INSTITUTE

Program Schedule

8:30-9:00	Check-in and continental breakfast
9:00-9:05	Introduction-**Mike Sadowsky,** MPGI Co-Director
9:05-9:25	**Brian Piasecki** *(PI and Advisor-Carolyn Silflow)* *"Centriolar maturation and ciliary formation in Chlamydomonas"*
9:25-9:45	**Yadong Huang** *(PI and Advisor-Sue Gibson)* *"Identification of a protein kinase and a RING-finger protein involved in plant sugar response"*
9:45-10:05	**Carrie Ketel**, PhD *(PI-Judith Berman)* *"Formation and movement of neocentromeres in Candida albicans"*
10:05-10:25	**Junfeng Gao** *(PIs and Advisors-Lynda Ellis & Larry Wackett)* *"Improving infrastructure for pathway prediction"*
10:25-10:45	Break
10:45-11:05	**Anke Reinders**, PhD *(PI-John Ward)* *"A role for sucrose transport in sucrose signaling and pollen germination"*
11:05-11:55	**Daniel Voytas**, PhD, Professor, GCD, and Director, Arnold & Mabel Beckman Center for Transposon Research *"Precise engineering of plant genomes using zinc finger nucleases"*
11:55-1:15	Lunch and Poster Session

1:15-1:35	**Gaurav Pandey** *(PI and Advisor-Vipin Kumar)* *"Association analysis-based transformations for protein interaction networks: A function prediction case study"*
1:35-1:55	**Yung-Tsi Bolon**, PhD *(PI-Carroll Vance)* *"Gene expression profiling of soybean near-isogenic lines contrasting in seed protein and oil"*
1:55-2:15	**Arvind Bharti**, PhD *(PI-Nevin Young)* *"Insights into the corn genome"*
2:15-2:35	**Jun Li**, PhD *(PI-Ken Vernick)* *"Coding domain sequence prediction and alternative splicing detection in human malaria vector Gambiae"*
2:35-2:55	**William J. Haun** *(PI and Advisor-Nathan Springer)* *"Genetic and epigenetic control of imprinting at the Mez1 locus in maize"*
2:55-3:15	Break
3:15-3:35	**Erik Lysøe**, PhD *(PI-H. Corby Kistler)* *"The transcription factor, FgStuA, influences spore development, pathogenicity and secondary metabolism in the plant pathogenic fungus Fusarium graminearum"*
3:35-4:25	**Adrian Hegeman**, PhD, Asst. Professor, Horticultural Science *"Stable-isotope assisted metabolomics: Enhanced elemental composition assignments from accurate mass measurements"*
4:25-4:30	Closing remarks: **Neil Olszewski**, MPGI Co-Director

Centriolar maturation and ciliary formation in *Chlamydomonas*

Brian P. Piasecki and Carolyn D. Silflow

Speaker: Brian Piasecki
PhD Candidate, Plant Biology

Abstract:
The unicellular green alga *Chlamydomonas reinhardtii* is typically biflagellate, but forward genetic screens have identified "uniflagellar" mutants. All uniflagellar mutants (*uni1*, *uni2*, and *uni3*) characterized thus far contain ultrastructural defects in the basal body or in the transition zone between the basal body and axoneme. Further, all uniflagellar mutants preferentially assemble a flagellum from the older basal body positioned *trans* to the eyespot. Thus, uniflagellar mutants allow for a molecular dissection of the pathways leading to flagellar assembly and basal body maturation. The *UNI2* gene encodes a novel coiled-coil protein with a potential homologue in the human genome. We rescued the *uni2* mutant phenotype with an HA-epitope tagged gene construct. Immunoblot analysis demonstrated that the Uni2 protein migrates as at least two molecular-weight variants that can be converted to a single form with phosphatase treatment. Synthesis of Uni2 protein is induced during cell division cycles; accumulation of the phophorylated form coincides with assembly of transition zones and flagella at the end of the division cycle. Using the Uni2 protein as a cell cycle marker of basal bodies, we observed the sequential assembly of new probasal bodies beginning at prophase. To test for a genetic interaction of the *UNI2* gene with other *UNI* genes, we constructed double mutants of *uni2* with both *uni1* and *uni3*. Ultrastructural analysis of serial cross sections through *uni1* and *uni2* single and double mutant cells demonstrated that they are deficient in triplet to doublet microtubule "transition" at the distal end of the basal body. These mutants provide the first mechanistic insights into the pathway for transition of triplet to doublet microtubules during ciliary assembly and suggest that the pathways mediating microtubule transition and basal body maturation overlap.

Identification of a protein kinase and a RING-finger protein involved in plant sugar response

Yadong Huang, Chun-Yao Li and Susan I. Gibson

Speaker: Yadong Huang
PhD Candidate, Plant Biology

Abstract:
In plants, metabolism and development are tightly integrated. For example, the levels of soluble sugars, such as glucose and sucrose, regulate diverse metabolic and developmental pathways. To elucidate the role of soluble sugars in plant metabolism and development, our lab has identified and characterized several "sugar insensitive" (*sis*) mutants of Arabidopsis. These *sis* mutants are resistant to the inhibitory effects of high concentrations of sucrose and glucose on early seedling development. Recently the genes corresponding to two *sis* mutants, *sis3* and *sis8*, were identified by map-based cloning approach. The *SIS3* gene encodes a protein that is predicted to be a member of the RING-finger family and contains transmembrane domains. The *SIS8* gene is predicted to encode a protein kinase. Current efforts are focused on characterizing SIS3 and SIS8. *In vitro* E3 ligase assays have confirmed the polyubiquitination function of *SIS3*. The subcellular locations of both SIS3 and SIS8 have been determined via transient expression assays. Possible interacting partners of *SIS8* were identified through a yeast two-hybrid screening. *35S-SIS3* and *35S-SIS8* plant were generated and their seed germination and early seedling development response to high sugar levels and to abscisic acid and gibberellins were assayed. The expression of sugar-regulated genes in the mutant and overexpressor lines of *sis3* and *sis8* have been studied by Q-PCR. Possible signaling mechanisms of *SIS3* and *SIS8* are discussed.

Formation and movement of neocentromeres in *Candida albicans*

Carrie Ketel, Mark McClellan, Helen Wang, Kelly Bouchonville and Judith Berman

Speaker: Carrie Ketel, PhD
PD Fellow, GCD

Abstract:
C. albicans has small (3-4.5 kb) regional centromeres (*CENs)* defined by a core region that binds CEN-H3; unlike other regional *CENs,* the core is not flanked by pericentric heterochromatin. The only sequence shared on two isochromosomes [i(5L)] (which confers fluconazole resistance) and i(5R) (found in some MTLhom strains)] is an inverted repeat (IR) flanking a 'central core' (CC) sequence that binds CENH3/Cse4p, the centromere-specific histone H3 isomer. This *CC+IR* organization resembles the DNA organization at regional centromeres in other organisms, including *S. pombe* and humans. We asked if the entire *CEN5 (CC5+IR)* region was necessary for centromere function by replacing it with *URA3*. Transformants were obtained at a 10X lower frequently than controls and the frequency of correct transformants were 5X lower but fluctuation analysis of the Δ*CEN5::URA3* strains detected ~wt levels of chromosome stability, implying that a neocentromere (functional centromere associated with new DNA) formed elsewhere on the chromosome. Interestingly, selection on 5-FOA, which kills cells expressing *URA3*, gave rise to Δ*CEN5::URA3* strains that formed colonies on both 5-FOA and SD-Uri and the FOA^R strains could subsequently grow on SD-Uri and vice versa. Southern blot analysis and RT-PCR data support the idea that the *URA3* marker that replaced *CEN5* is present on Chr5 in FOA and –Uri grown cells and is reversibly silenced. Chromatin-immunoprecipitation with antibody to Cse4/CEN-H3 in Δ*CEN5::URA3* strains revealed that the position of the *neoCEN* was dependent upon the expression status of *URA3:* Cse4/CEN-H3 was associated with (and presumably caused the silencing of) the *URA3* DNA in cells growing on 5-FOA; Cse4/CEN-H3 was associated with DNA adjacent to *URA3* in cells growing on SD-uri. In summary, *C. albicans* has regional centromeres that can form neocentromeres upon deletion of a *CC+IR* structure. Furthermore, the neoCEN can transiently associate with, and thereby silence, different DNA sequences, depending upon the selective growth conditions. Thus, *C. albicans* provides a very simple model for studying the function of regional centromeres and the positioning of neocentromeres.

Improving infrastructure for pathway prediction

Speaker: Junfeng Gao
Research Assistant, BTI

Abstract:
The UM-BBD Pathway Prediction System (UM-PPS, http://umbbd.msi.umn.edu/predict/) predicts microbial catabolism of organic compounds. We improved UM-PPS infrastructure to improve pathway prediction results. We added the ability to allow relative reasoning and variable aerobic likelihood. One relative reasoning entry decreased choices 75% with no loss of sensitivity. Variable aerobic likelihood gives more accurate likelihood for rules triggered by substrates with certain chemical structures. Predictions are improved.

A role for sucrose transport in sucrose signaling and pollen germination

Speaker: Anke Reinders, PhD
Research Associate, Plant Biology

Abstract:
Sucrose transporters in plants play a pivotal role in carbon partitioning between source and sink tissues. As the main transport form of fixed carbon in most plants, sucrose has to be actively loaded into the transport network, the phloem. Plants with

decreased activity of the phloem-loading sucrose transporters are severely affected in their growth and development. In the model plant Arabidopsis sucrose transporters belong to a small gene family of 7-9 members called SUTs or SUCs. Only one of these transporters, AtSUC2, is required for phloem loading. Others members of the family are expressed in specific plant parts but their role has been less well characterized. AtSUC1, a close homolog of AtSUC2, is highly expressed in flowers, especially in pollen, in roots and in trichomes. Its transport activity, determine by two-electrode voltage clamping, is very similar to that of AtSUC2. Here we show that loss of AtSUC1 leads to a defect in pollen function in vivo as well as low pollen germination in vitro. In addition, atsuc1 mutants were lacking in the induction of anthocyanin synthesis in response to sucrose and were generally affected in the expression of many genes involved in anthocyanin biosynthesis. Our results indicate that AtSUC1 has an important role in both reproductive and vegetative tissues.

Precise engineering of plant genomes using zinc finger nucleases

Speaker: Daniel Voytas, PhD
Professor, GCD, and Director, Beckman Center for Transposon Research

Abstract:
Engineered zinc finger nucleases induce targeted genome modifications with high efficiency in many different cell types. ZFNs consist of a Cys_2His_2 zinc finger domain engineered to bind a particular DNA sequence and the non-specific nuclease domain of the FokI restriction enzyme. ZFNs introduce double-stranded breaks at specific DNA sequences and thereby vastly increase the rate of homologous recombination at the cleaved locus. Targeted mutations can also be introduced into specific loci through imprecise repair of the broken chromosome by non-homologous end-joining. To fully capitalize on the potential fo ZFNs for genome modification, we implemented a rapid and robust platform

for constructing zinc finger arrays. Using this platform, we engineered numerous ZFNs that recognize a variety of endogenous vertebrate and plant genes. In plants, the tobacco *SuRB* locus is being used as a model to establish key parameters for high frequency recombination. *SuRB* encodes acetolactate synthase, an enzyme that carries out the first step in branched chain amino acid synthesis. We have shown that ZFN-induced recombination can introduce mutations throughout *SuRB* that confer resistance to various herbicides. The precise genome changes afforded by ZFN-mediated gene targeting can help discern the function of the many genes revealed through the various genome sequencing projects. In addition, ZFNs can be used to develop new crop varieties, including those that better withstand pests, have enhanced food value, and produce compounds of industrial importance. Because gene targeting introduces changes in plant genomes in a highly specific and controlled manner, crops generated through recombination may be met with greater public acceptance than traditional genetically modified crops.

Association analysis-based transformations for protein interaction networks: A function prediction case study

Speaker: Gaurav Pandey
Research Assistant, Computer Science & Engineering

Abstract:
Protein interaction networks are one of the most promising types of biological data for the discovery of functional modules and the prediction of individual protein functions. However, it is known that these networks are both incomplete and inaccurate, i.e., they have spurious edges and lack biologically valid edges. One way to handle this problem is by transforming the original interaction graph into new graphs that remove spurious edges, add biologically valid ones, and assign reliability scores to the edges constituting the final network. We investigate currently

existing methods, as well as propose a robust association analysis-based method for this task. This method is based on the concept of h-confidence, which is a measure that can be used to extract groups of objects having high similarity with each other. Experimental evaluation on several protein interaction data sets show that hyperclique-based transformations enhance the performance of standard function prediction algorithms significantly, and thus have merit.

Gene expression profiling of soybean near-isogenic lines contrasting in seed protein and oil

Speaker: Yung-Tsi Bolon, PhD
Research Associate, Agronomy & Plant Genetics

Abstract:
Soybean profitability is affected by protein and oil content. Thus, an understanding of the genetic controls on protein and oil yield is important for future soybean improvement. In this study, we used Affymetrix soybean genome arrays with >37,500 Glycine max probe sets to compare gene expression profiles from nearly identical soybean lines that differ in protein and oil. Gene expression profiles were obtained from four different stages of the developing soybean seed. Analysis of differential gene expression produced a significant list of less than 200 genes. Presumably, these variations are related to the difference in protein and oil content between these two lines. These results demonstrate the power of gene expression analysis to contrast near-isogenic lines. Further investigation may provide new insight into the genes and pathways involved in protein and oil accumulation in the soybean seed.

Insights into the corn genome

Speaker: Arvind Bharti, PhD
Research Associate, Plant Pathology

Abstract:
The two major clades of the grass family (*Poaceae*) have been estimated to have diverged ~80 mya (million years ago). Rice belongs to the BEP clade while maize and sorghum belong to the PACCAD clade. While rice and sorghum are diploids, maize is known to be an ancient allo-tetraploid with chromosomes composed of many homeologous segments of the two parental genomes that subsequently underwent loss of one copy for over half of its duplicated genes. Sequencing a sample of these homeologous regions from different maize chromosomes and comparing the linked genes with orthologous regions from rice and sorghum has permitted the alignment of genes derived from common ancestral grass chromosomes. Two large regions of *Zea mays* ssp. *mays* cv B73 (8 Mb from chromosome Zm1S and 7 Mb from Zm9L) have been sequenced and compared in their gene content and architecture [Genome Res.16:1241. 2006]. These regions are orthologous to a segment of rice Os3S (4.5 Mb) and to sorghum Sb1L (6.7 Mb). Alignments of 36 gene models, as representatives of orthologs common across all four genomic regions, have been used to estimate divergence times of the two progenitors of maize. Our results support a tetraploid origin of maize with an estimated divergence time for the two maize progenitors to be around 30 Mya. Our data also confirms that the two maize progenitors and sorghum diverged almost at the same time and that any one of the maize progenitor is not more closely related to sorghum than to the other maize progenitor. We have also observed that the retention of an ortholog on any one of the duplicated region in maize is random.

Coding domain sequence prediction and alternative splicing detection in human malaria vector *Gambiae*

Speaker: Jun Li, PhD
Senior Research Associate, Microbiology

Abstract:
Genome-wide research needs good coding domain sequence (CDS) structure prediction and alternative splicing information. Current Anopheles gambiae CDS prediction (ensembl) based on comparative algorithm is incomplete with ~40% CDS lacking start and/or stop codon. In addition, genome-wide alternative splicing has not been reported in Anopheles gambiae. We collected 819 Anopheles gambiae gene structures from experimental data, and trained the ab initio gene prediction model of GlimmerHMM with this set. This trained GlimmerHMM predicted 14916 CDS with sensitivity of 95% and specificity of 90%. Comparison of ab initio predicted CDS set to ensemble CDS showed that 92% CDS is overlapping. To take advantage of the two major gene prediction algorithms, we synthesized the ensembl and ab initio predicted coding sequences into a single set using exon-gene-union algorithm followed by open-reading-frame-selection algorithm. 16191 CDS were created. Among them, 95% CDS are complete. In the combinational CDS, 8510 are identical to ensembl, 3999 CDS were extended, 2522 experienced major structure change, and 1160 are novel. We used ~165K EST to assess the quality of predicted CDS. Results indicated that 61.4% of combinational CDS were supported by these EST. Same set of EST was also used to detect the alternative splicing, and 1512 genes showed alternative splicing. The alternative splicing pattern of A. gambiae is more like plants than mammals. This genome annotation software has been integrated into our open-source package – "SweetGenomics".

Genetic and epigenetic control of imprinting at the *Mez1* locus in maize

Speaker: William J. Haun
PhD Candidate, Plant Biology

Abstract:
Genomic imprinting is the differential expression of a gene based on its parent-of-origin. Imprinted genes play a critical role during early development in both animals and plants. In *Arabidopsis* and maize, over a dozen imprinted genes have been identified and a detailed model of the mechanism of imprinting has been formulated. However, gaps in our knowledge regarding the 'how' and 'why' of imprinting still remain. We have characterized the epigenetic marks present on the active maternal allele and silent paternal allele of maize imprinted gene *Mez1*. The two parental alleles exhibit differences in both DNA methylation and histone modification patterns in the 5' proximal region. We also are using transposon insertion alleles in the 5'*cis*-proximal region of M*ez1* to study the genetic requirements for imprinting. These transposons insertions disrupt imprinting in unexpected ways and may help to shed light on the mechanisms of imprinting. In addition, we are able to study the biological role of imprinting by comparing the phenotype of kernels with and without imprinted expression of *Mez1*.

The transcription factor, *FgStuA,* influences spore development, pathogenicity and secondary metabolism in the plant pathogenic fungus, *Fusarium graminearum*

Erik Lysøe, Matias Pasquali, Andrew Breakspear, Sonja S. Klemsdal, and H. Corby Kistler

Speaker: Erik Lysøe, Ph.D.
PD Associate, Plant Pathology

Abstract:
The filamentous fungus *Fusarium graminearum* causes extensive losses on cereals world-wide and contaminates harvested grain

with mycotoxins, whose levels in the food supply are strictly regulated. We deleted the *FgStuA* gene in *Fusarium graminearum* and demonstrate its involvement in several different processes, such as spore development, pathogenicity and secondary metabolism. The *FgStuA* protein is a member of the APSES family which regulates morphogenesis and virulence in ascomycetes. *FgStuA* is closely related to *FoStuA* in *F. oxysporum* and *StuA* in *Aspergillus*, but unlike *FoStuA* mutants, the *FgStuA* mutants were greatly reduced in pathogenicity both on wheat and apple slices. Reduced pathogenicity may be due to decreased levels of trichothecene mycotoxins (<1% the levels of wildtype), which have shown to be pathogencity factors in wheat. Microarray analysis of the *FgStuA* mutant during plant infection shows that the trichothecene biosynthetic pathway is completely turned off. *FgStuA* mutants also were greatly reduced in asexual sporulation and produced no perithecia. The asexual spores produced also showed a slower germination than the wild-type, as well as reduced growth on agar plates. Microarray analysis during "sporulation conditions", (when the wildtype produced asexual spores and the mutant produced no spores), showed that genes encoding several groups of cell-wall related proteins such as chitinases, glycanases and GPIs, mostly were down-regulated in the mutant. Other putative spore-related genes, such as hydrophobins, also were highly down-regulated in the mutant. Genes found in the MIPS functional categories Transcription, Protein synthesis and Proteins with binding functions were extremely up-regulated in the *FgStuA* mutant during sporulation conditions, similar to wildtype developmental time-points during conidia- and ascospore germination. On V8 agar, the *FgStuA* mutant has a white phenotype compared to the red wild-type. Under sporulation conditions, 17 contiguous genes, including all known genes for biosynthesis of the red pigment aurofusarin, were virtually turned off in the mutant. Also, eight contiguous genes encoding a putative PKS-NPS hybrid were found to be down-regulated. We conclude that the *FgStuA* protein in *F. graminearum* functions as a global regulator conferring cross pathway control of sporulation, pathogenicity and gene clusters important for secondary metabolism.

Stable-isotope assisted metabolomics: Enhanced elemental composition assignments from accurate mass measurements

Speaker: Adrian Hegeman, PhD
Assistant Professor, Horticultural Science

Abstract:
The high degree of molecular diversity inherent in small molecule populations derived from biological samples poses significant analytical challenges. Metabolomics is an emerging field, which seeks to define the identity, quantity and location of all small molecule components of a biological system. This is, to date, not possible even for very simple systems. Various analytical methodologies are being employed for metabolomics including NMR, FT-IR, GC-MS, CE-MS and LC-MS, and each has its own strengths and weaknesses. Although GC-MS approaches are currently the most accessible and reliable for compound ID and quantification they are limited by requirements for volatile and thermally stable analytes. LC-MS has the potential to be much more broadly applicable than GC-MS, but is currently restricted by the lack of reliable compound identification tools and inconsistencies in quantification caused by matrix effects. An important element of my research program concerns the development of new metabolomics methodologies. We are currently exploring the application of ^{13}C and ^{15}N stable-isotopic metabolic labeling of plants to provide additional analytical information from LC-TOF analyses of extracted metabolites. We have devised an approach to constrain the numbers of formulae derived in elemental composition assignments from accurate mass measurements in a semi-automated manner. These assignments constitute the first step in the development of a high throughput process for the relative quantification of metabolites via ^{13}C-metabolic labeling. This methodology promises to help mitigate the negative analytical consequences of matrix effects, and provide a means for internal control of extraction fractionation during sample preparation.