

From Genes to Social Behavior in the Honey Bee

Social behavior in the honey bee is exquisitely sensitive to age, environment and genotype. Microarray analysis may identify the genes by which these factors act on behavior.

Understanding the relationship between genes and behavior is a fundamental challenge in biology. Behavior is a function of both “nature” and “nurture.” However, very little is known about how nature (i.e., inherited genetic factors) and nurture (i.e., social and other environmental factors) influence behavior. Which genes actually contribute to genetic variation in behavior? Are these genes expressed differently or do they produce different proteins? How does the environment mediate changes in behavior? Can the environment act on gene expression in the brain? A third

important component in behavior is time: as individuals age and pass through different stages of their lives, their behavioral responses to stimuli change in predictable ways. Does regulation of gene expression in the brain play a role in age-related changes in behavior?

The honey bee, *Apis mellifera*, is a highly social animal that exhibits profoundly interesting behavioral phenomena, such as division of labor, kin recognition, and communication via the dance language.

continued on page 2



Kenneth Lorenzen

Update on the Biotechnology Initiative

Planning for the Post Genomics Institute (PGI) is continuing with \$7.5 million in planning funds provided from the State of Illinois as part of Governor Ryan’s VentureTECH program. An Architecture/Engineering firm has been hired and is working with a group of faculty and staff to develop a detailed program statement for the ~110,000 ft² building (estimated cost = \$75 million). The Post Genomics Institute will contain laboratory space for research groups organized around scientific themes, teaching laboratories and facilities for the Biotechnology Center. Construction will begin in 2001 assuming the State of Illinois provides the remaining funds as expected. A committee is being formed to conduct a national search for the PGI Director; nominations for this position can be submitted to Tony Waldrop who will chair the search committee.

In addition to the PGI planning monies, the campus has received \$1.6 million recurring to hire faculty and staff associated with the programmatic themes of the Post Genomics Institute. An additional \$3.2 million recurring is expected from the State of Illinois in the next fiscal year. The Provost and Vice Chancellor for Research have announced that the first \$1.6 million will be released to units in a campus competition to provide new faculty lines. Proposals for these funds are to be submitted by November 20 with anticipated announcement of those chosen to be announced in December. It is expected that this process will be one of the steps in defining the foci for the Post Genomics Institute. To support these new hires, campus will provide \$4 million in renovation funds for laboratory space and offices and additional funding for startup packages. Departments and colleges are expected to contribute funding as well.

—Tony Waldrop, Vice Chancellor for Research

Winter 2001

IN THIS ISSUE

2

Data Management at the W.M. Keck Center Bioinformatics Unit

4

Molecular Subtraction: A Tool for Large-Scale cDNA Sequencing Projects

4

Tools for Proteome Analysis

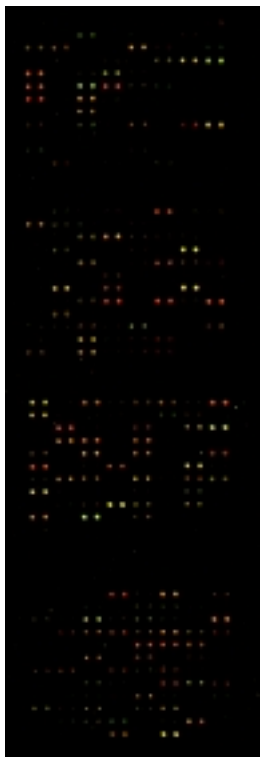
6

Faces in the News

7

Facility Updates

From Genes to Social Behavior in the Honey Bee, continued



Pilot honey bee array representing approximately 350 ESTs. About half of the ESTs are random EST clones and the other half were selected based on bioinformatic analysis (these are predicted to represent genes involved in interesting neurobiological, behavioral or other processes). Messenger RNA was extracted from the heads of bees collected from the inside of a bee hive (labeled red) and from the heads of bees that were returning to the hive (labeled green). Several of the intense red spots represent genes that encode royal jelly proteins, known to be produced by nurse bees but not forager bees. Spots higher in the green channel include a glutamate transporter (known to be higher in the brain of foragers) and several genes involved in glutamate synthesis.

Bee behavior is related to age, genetics and environment. A worker bee begins her adult life by working inside the hive, changing from cell cleaning to brood care (“nursing”) and then to food storage activities as she ages. After two to three weeks inside the hive, she then shifts to collecting nectar and pollen (“foraging”) for the remainder of her life. The speed with which a bee matures into a forager varies greatly. Some bees show precocious behavioral development and become foragers at relatively younger ages, while others develop more slowly and do not initiate foraging until later in life. This variation is due to both genetic and environmental factors. Bees of different genotypes show differences in rate of behavioral development. As a highly social animal, the bee’s pace of behavioral maturation also is influenced by interactions with other colony members. Bees can accelerate, delay, and even reverse their behavioral ontogeny in response to changing conditions.

How can we understand the bee’s behavioral development at the molecular level? Under the leadership of **Prof. Gene Robinson**, a Professor in the Department of Entomology and Director of the Bee Research Facility, our lab initiated a large gene discovery project focused specifically on the honey bee brain. We generated a bee brain Expressed Sequence Tag (EST) resource by partially sequencing the 5’ end of 20,352 clones isolated from a normalized cDNA library from dissected

bee brains. The library was constructed by Prof. Bento Soares, University of Iowa, a pioneer in the normalization technique, using mRNA collected from bees at the Bee Research Facility. These clones were sequenced by the High Throughput Sequencing and Genotyping Unit at the Keck Center. After sequencing of approximately 10,000 clones, a molecular subtraction was performed, and sequencing of the subtracted library was continued. Sequences were screened for a variety of criteria, resulting in 14,632 high-quality sequences of average read-length 593 bp. This read-length is exceptionally high for an EST project, and has no doubt contributed greatly to our bioinformatics efforts. The high-quality sequences were assembled by the Bioinformatics Unit (also at the Keck Center), resulting in 9195 total unique sequences. Approximately 40% of these sequences are highly similar to gene sequences from other organisms. In close collaboration with the Bioinformatics Unit, our lab has initiated annotation of these sequences. Based on high similarity with *Drosophila* genes alone, more than 2200 bee ESTs can be assigned tentative molecular function.

What can we do with these ESTs? In addition to greatly accelerating the rate at which individual genes can be isolated and studied, having so many ESTs in hand allows a functional genomics approach. All 9195 unique bee sequence clones can be spotted on a single glass slide (microarray) us-

Data Management at the W.M. Keck Center Bioinformatics Unit

With the advancement of various genome projects, biologists have gained access to complex databases that are orders of magnitude larger than ten years ago. Gigabytes of data must be converted to a usable form, transforming complex information into insight. At present, without the right tools, new discoveries are being hampered by information overload. Two major areas must be addressed in the field of bioinformatics: 1) the creation, management and accession of large repositories of sequence, structural and functional data; and 2) the creation of appropriate software tools that assist the scientist in analyzing the data. Convenient management of projects and efficient processing of data are urgent needs for biologists doing large-scale

genome research. Using Java Servlets, a powerful and portable CGI-like technology, we have developed an online project management system, called the Genome Project Management System (GPMS) that integrates databases with a suite of bioinformatics tools.

The Genome Project Management System integrates a data pipeline that links various analysis programs and DNA sequence data together to ensure smooth data flow. Data generated at DNA sequencers at the Keck Center are automatically parsed into an Oracle database (SequenceDB), then edited by running various analysis programs, and ultimately put into the hands of the users. In addition, GPMS has integrated several

	Query	Query Length (nt)	HSP	e-value (aa)	Subject Length (aa)	Align. Length (aa)	Identifies	Positives	EBlog	Gene Symbol	Gene Name	Molecular Function	Biological Process
1	BB270010B10A12.F	742	96	2E-90	1012	73	56%	71%	EBlog0034432	EG-90H7.7	%N-methyl-D-aspartate selective glutamate receptor		
2	Contig2738	704	353	6E-88	456	195	87%	92%	EBlog0034563	GluClAgg;	%glutamate-gated chloride channel	%ion transport	
3	BB270009A10D3.F	707	278	2E-78	496	163	84%	88%	EBlog0035240	LochD	Ligand-gated chloride channel homolog 3	%gamma-aminobutyric acid-inhibited chloride channel	%ion transport
4	BB260004A10F4.F	768	336	6E-80	827	267	63%	79%	EBlog0035402	CG12682	%ubiquitin isopeptidase T	<deubiquitylation	
5	BB270026B10B6.F	720	81	6E-16	1426	159	36%	48%	EBlog0035731	Egfr	Epidermal growth factor receptor	%surkin receptor	<dorsal/ventral axis determination
6	Contig726	746	280	4E-76	159	152	54%	97%	EBlog0036416	Vha1G	%vacuolar H+ pump<->ATPase 16kD subunit	%hydrogen-transporting hetero-sector ATPase	

Examples of annotated bee EST sequences resulting from BLASTX similarity searches with Drosophila predicted genes. More than 2200 bee sequences can be assigned tentative molecular function based on functional assignment in Drosophila.

ing a robot at the Functional Genomics Unit. These microarrays allow measurement of gene expression for each gene on the array. Whole bee heads have been assayed for gene expression using a small (~350 spot) pilot array (see Figure). These studies indicate that large differences are present between bees collected from the inside of the hive and bees collected that were returning to the hive (likely reflecting major differences between “nurses” from inside and “foragers” from outside). Our studies also indicate that we can amplify small amounts of messenger RNA from as little as a single bee brain for use in microarray studies. This will allow us to study gene expression levels on a genomic scale in individual bee brains.

The close association between bees and humans for millennia has resulted in a great wealth of knowledge in the natural social life of the honey

bee. This allows us to precisely manipulate the conditions in the colony that are known to regulate behavior. By combining the powerful methods employed in bee behavioral studies with functional genomics, we can begin to address the question: is variation in gene expression in the brain a mechanism that underlies behavioral diversity in the bee colony?

—Charles Whitfield

This project was supported by grants from the University of Illinois Critical Research Initiative and a Burroughs-Welcome Innovation Award in Functional Genomics to GER and a National Science Foundation Postdoctoral Fellowship in Biological Informatics to CW.

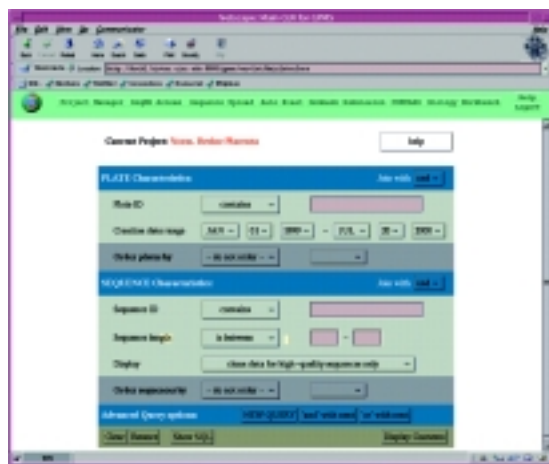


Figure 1. SeqDB Access module: from this module users can query and retrieve the sequences for a particular project.

bioinformatics tools. The SequenceDB Access program (Fig. 1) allows the user to query his/her own data from SequenceDB through a user-friendly interface. From the interface, the user can browse the sequences directly on the Web and can select any number of sequences to download or to pipe into other modules for further analysis. The Upload Sequence program lets the user upload sequences from the user’s local machine to the GPMS server system for analysis (Fig. 2). The AutoBlast program provides a batch high throughput BLAST search (Fig. 3). The search results are then parsed and deposited into another Oracle database. The user can retrieve and analyze the

continued on page 8



Molecular Subtraction: A Tool for Large-Scale cDNA Sequencing Projects

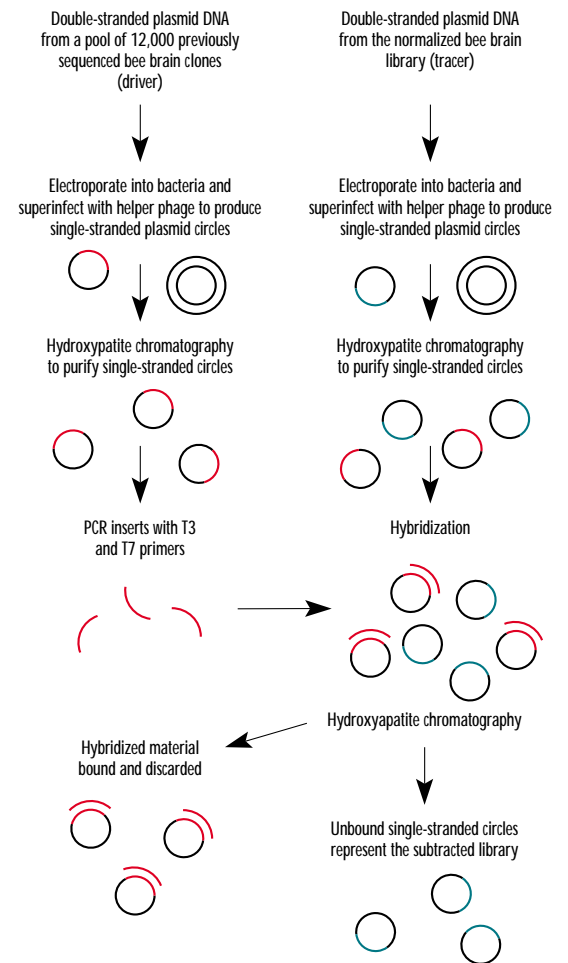
Gene discovery through mass sequencing is critically dependent on the use of cDNA libraries. The latter are generated by reverse transcription of cell-specific mRNA populations and are usually cloned in plasmid. However, the power and cost-effectiveness of large-scale sequencing is hindered by the presence, in all cells, of a class of mRNAs whose mass makes up to 65% of total message, yet represents at most only 2,000 distinct species.

One way to avoid the repetitive, and expensive, sequencing of highly redundant cDNAs is to remove them from a cDNA library. This is accomplished using a technique generically known as molecular subtraction. One version of this approach involves the hybridization of a single-stranded cDNA library (the “tracer”) with the collection of PCR-amplified cDNAs one wishes to eliminate (the “driver”) (Bonaldo et al. 1996). Following hybridization, duplexes are separated from single-stranded molecules on hydroxyapatite to generate the subtracted library (see figure).

A subtracted library is an extremely valuable reagent – it makes possible the efficient identification of novel genes as well as the economical sampling of the more rare mRNAs present in a particular cell type. **Jose R. Pardinás, Ph.D. and Rachel Schwartz** at the W.M. Keck for Comparative and Functional Genomics have applied this methodology, with great success, to large-scale sequencing of bovine placenta and bee brain cDNA libraries.

—Edited by B. Whitmarsh

Reference: Bonaldo, M.F., G. Lennon, and M.B. Soares. 1996. Normalization and subtraction: Two approaches to facilitate gene discovery. *Genome Res.* 6: 791-806.



Tools for Proteome Analysis

Moving towards proteomics, the Protein Sciences Facility developed 2D gel electrophoresis in addition to the existing mass spectrometry and sequence analysis services. 2D gel electrophoresis is a useful tool for comparing protein expression of cells and/or tissue in two different biological states.

This multi-step experiment begins with proper sample preparation. A mixture of proteins isolated

from cell extract is denatured in a lysis solution to achieve complete solubilization of all proteins. Urea is used for complete denaturation and a strongly reducing agent is added to break disulfide bonds and keep all proteins in a reduced state. In the first step of the gel electrophoresis, the protein mixture is separated using isoelectric focusing. A comprehensive range of overlapping

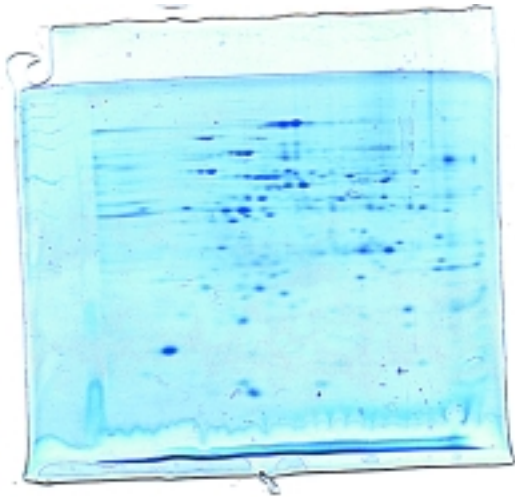


Figure 1: 2 D gel of a bacterial protein extract. First dimension: IPG Dry Strip pH 4-7; second dimension: gradient SDS PAGE gel (4-20%); stain: Coomassie Blue.

DryStrip Gels with immobilized pH gradients covering narrow (1 pH unit), medium (3-5 pH units) and wide (7 pH units) pH ranges, can be tested for optimization. After separation by isoelectric focusing, a second dimension electrophoresis is run, based on separating proteins by molecular weight. A successful separation results in a well focused, highly resolved protein pattern (figure 1) displaying up to hundreds of proteins. The success mostly depends on accurate sample preparation. By careful analysis of the intensity and/or the absence or presence of spots, the investigator can search for key proteins expressed under certain conditions.

Further analysis of the selected proteins may lead to protein identification. Proteins of interest are excised from the gel to undergo in-gel digestion. The digest is subjected to separation by capillary HPLC to separate the peptide mixture. Isolated peptides can be subjected to N-terminal Edman Sequencing (figure 2) to elucidate the amino acid sequence. Sequence analysis of several peptides provides the key information to identify the protein.

Another route can also lead to protein identification. The peptide mixture can be analyzed directly by using MALDI-TOF mass spectrometry yielding a characteristic peptide map (figure 3). Molecular weight analysis of the identified peptides is in many cases accurate enough for positive protein identification via a database search. Partial se-

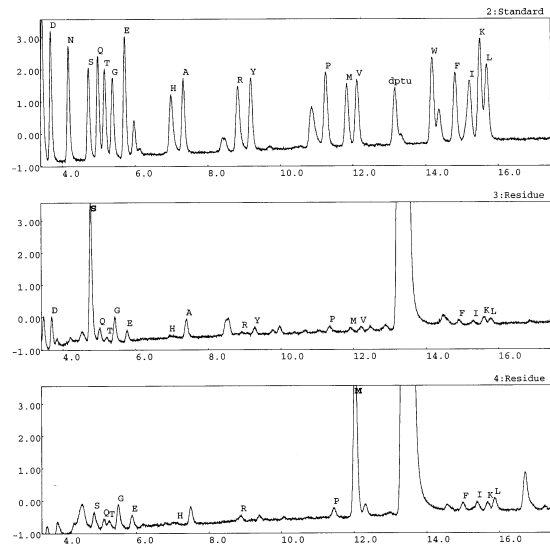


Figure 2: Sequence analysis chromatograms showing the amino acid standard and 2 amino acid residues of a sample.

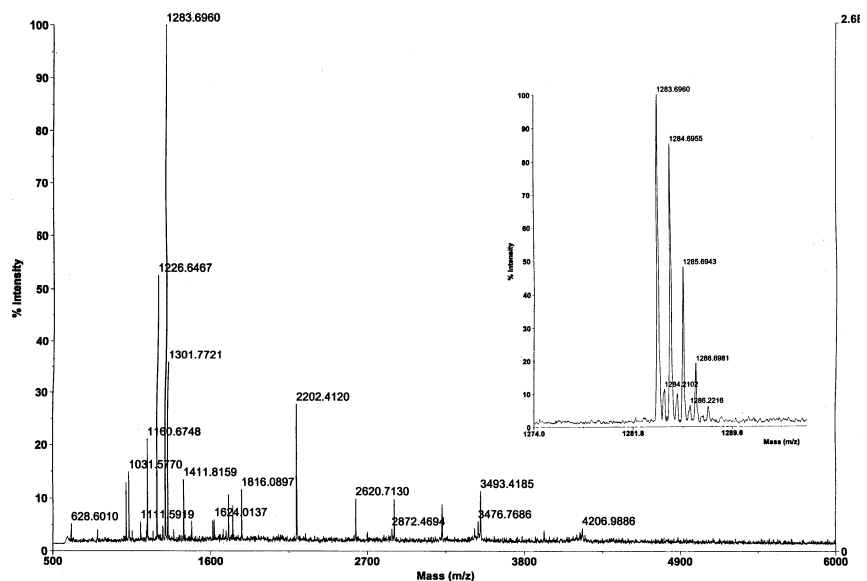


Figure 3: MALDI mass spectrum of a peptide mass map obtained from tryptic in-gel digest.

quence information can also be obtained by performing MALDI post-source decay analysis.

The data displayed have been generated at the Protein Sciences Facility. Capillary HPLC separation will be available in January of 2001. For further information and project consultation contact the facility director Dr. Henriette Remmer (remmer@uiuc.edu, phone: 333-4695).

—B. Whitmarsh



Kathleen Brinkmann,
*Director, Biotechnology
Center Placement Office*

New Director for the Biotechnology Center Placement Office

Please welcome Kathleen Brinkmann as our new Placement Director. The Biotechnology Center Placement Office provides a central location for advanced degree job candidates, including M.S.s, Ph.D.s and Ph.D.s currently in postdoctoral positions, for job placement in industry and academia. As a campus-wide office, it serves advanced degree students from 19 departments in six different schools and colleges who pursue research in a biological area.

The new Placement Director, Ms. Brinkmann, began her position on July 31, after receiving some initial training from Catherine Connor, who retired the end of July to pursue her new endeavors. Ms. Brinkmann has a strong science background, based on nine years of work in research laboratories here on campus. She also worked in industry and owned a business for several years, broadening her experience and gaining insight into the business community. Community outreach, which included working for the Orpheum Children's Science Museum, and teaching at Parkland College taught her organizational and interpersonal skills and are examples of her initiative and creativity. Ms. Brinkmann is energetic, a self-starter and brings enthusiasm to her new position. We wish her a good start and a productive first year.

Please encourage your graduate students and postdocs who are looking for a job to contact the Biotechnology Center Placement office at 333-1378 or via e-mail at brinkman@staff.uiuc.edu to set up an appointment with Ms. Brinkmann. Ms. Brinkmann will discuss job strategies with the student, including how to prepare an effective resume, what companies might be the best match for your student's career goals, how to prepare for an interview and other useful information. In addition, please ask them to contact the office and find out when recruiters from major companies will visit the campus to conduct interviews.

—B. Whitmarsh

Lei Liu Named NCSA Faculty Fellow

Dr. Lei Liu, Director of the Bioinformatics Unit at the W. M. Keck Center has been named one of eleven NCSA Faculty Fellows for the academic year 2000-2001. The Faculty Fellows program works to extend opportunities in advanced com-

puting and information technology to Illinois faculty. Faculty Fellows have access to NCSA's high-performance computers, visualization and virtual reality environments, and opportunities to collaborate with colleagues at NCSA and throughout the National Computational Science Alliance (<http://access.ncsa.uiuc.edu>). During the one-year program Dr. Liu's group will collaborate with scientists from the Data Visualization and Data Mining Division at NCSA to tackle the complex problems of information integration of genomic data. Currently we are working on automating annotations of unique EST sequences that we identified in our on-going cattle and honey bee large-scale EST sequencing projects. Once these annotations are available, we will use the information to analyze and interpret the results of microarray hybridization experiments for both species.

Bruce Chassy—Executive Associate Director for Outreach

Please join me in welcoming Professor Bruce Chassy, Assistant Dean College of ACES, as Executive Associate Director of the Biotechnology Center. Before accepting his new position, Prof. Chassy served as Head of the Department of Food Science and Human Nutrition and conducted research in molecular biology and biotechnology of lactic acid bacteria in food and dairy applications. In his new role for outreach and extension, Professor Chassy will communicate the nature and significance of biotechnology research and development at the University of Illinois at national and international conferences and symposia. His long-standing reputation as a distinguished scientist and academic leader make him ideally suited to serve as a spokesperson for the University of Illinois campus Biotechnology Initiative, including the role of the Post Genomics Institute and the Biotechnology Center.

Having served for many years as a member of professional societies, Prof. Chassy continues his involvement by interacting with many external groups at the national and regional level, such as the State legislature, Biotechnology Industry, policymakers such as the Food Advisory Council and the Institute of Food Technologists, major agricultural associations, and the public at large. To increase visibility of the strength of biotechnology research at the University of Illinois, Prof. Chassy is planning to create a biotechnology web site that will emphasize the major areas of biotechnology research conducted at this campus.

—B. Whitmarsh



**Bruce Chassy, Associate
Executive Director,
Biotechnology Center &
Assistant Dean,
Biotechnology and
Outreach, College of
ACES**

Facility Updates

Phosphor/Fluoroimager at the Flow Cytometry Facility



The Flow Cytometry Facility has a new imaging instrument from Molecular Dynamics, now a subsidiary of Amersham Pharmacia, operating and ready-for-use. The “**Typhoon**”, the latest upgrade from the “Storm”, combines multi-color fluorescence, chemifluorescence, and filmless autoradiography. The instrument has two excitation lasers, wavelength 532 nm and 633 nm, 7 filters and and 4 color scans. The range in sensitivity as a phosphorimager offers 5 orders of linear dynamic range. More detailed information is available at <http://www.mdyn.com/products/typhoon/default.htm>. Screen sizes are: 20 x 25 cm and 35 x 43 cm. To order screens please check this web site http://www.life.uiuc.edu/biotech/screens_cassettes.html.

The purchase of this instrument was made possible by contributions from the Illinois-Missouri Consortium, initiated by Prof. Torbert Rocheford, Department of Crop Sciences, the Biotechnology Center, funds from the Research Board for a proposal submitted by Prof. Mary Schuler, Department of Cell and Structural Biology, the School of Molecular and Cellular Biology, and Prof. Keith Kelly, Department of Animal Sciences.

Immunological Resource Center Updates



Bioreactor Services

The Immunological Resource Center (IRC) now offers Integra™ flask bioreactor chambers for in vitro production of monoclonal antibodies (for more information on this product visit <http://www.integrabiosciences.com/e-fcelline.html>). The reason for creating this service was to address new federal regulations that mandate in-vitro suitability testing of antibody producing cell lines prior to LACAC (Laboratory Animal Care Advisory Committee) approval of the classical ascites fluid method.

To the researcher who requires monoclonal, this mandate simply means that in order to have antibodies produced via the ascites method, in vitro techniques will have to be demonstrated as unsuitable for their research application; this will involve the actual testing of in-vitro produced antibody. The IRC staff will help the researcher in the evaluation of in-vitro options for each submitted request for antibody production. Antibody production estimates and pricing are in their formative stages, but the cost to set-up and successfully complete an Integra™ CL350 unit (~20 mg antibody) and \$500 (~60 mg antibody) for the Integra™ CL1000.

While this change in policy will raise the cost of antibody production, the antibody product is free of endogenous mouse immunoglobulins that are present in ascites produced antibody and the extension of this service brings the campus into full compliance with the new federal rules regulating the production of monoclonal antibodies via the ascites fluid method. For more information on this topic please visit the UIUC campus OLAR web site: <http://www.olar.uiuc.edu/Policies/celllines&ascites.htm>

New Fee Structures for Monoclonal and Polyclonal Antibody Development

As a direct result of transitioning from the ascites method to in-vitro production methods, more laboratory assistance is needed to develop and deliver the in-vitro bioreactor produced monoclonal antibody. In response to this need, Theresa Holly was hired and began her appointment in early October. To recover the costs associated with this additional appointment, a generalized service fee adjustment was calculated and applied to the IRC service fee. When preparing grant applications that involve antibody development and production, please contact Steven D. Miklasz, Director of IRC to calculate an appropriate budget.

—Steven Miklasz

Biotechnology Center at the University of Illinois at Urbana-Champaign

Harris Lewin, PhD • Director and Gutsell Professor of Animal Sciences

Robert Gennis, PhD • Executive Associate Director and Professor of Biochemistry

Bruce Chassy, PhD • Executive Associate Director for Outreach

Barbara Whitmarsh, PhD • Associate Director and Editor

Administrative Office • (217) 333-1695 phone; (217) 244-0466 (fax)

Please send comments on the newsletter, news, or questions to Barbara Whitmarsh at b-whitmarsh@uiuc.edu or contact the Administrative Office.

Data Management, continued from page 3

Figure 2. Sequence Browser in SeqDB Access module. This module allows users to perform several functions: view the progress of their sequencing project, view the quality of their sequences, download data to their local machine, export data to the AutoBlast server for running a BLAST search, or attach sequences for Genbank submission.

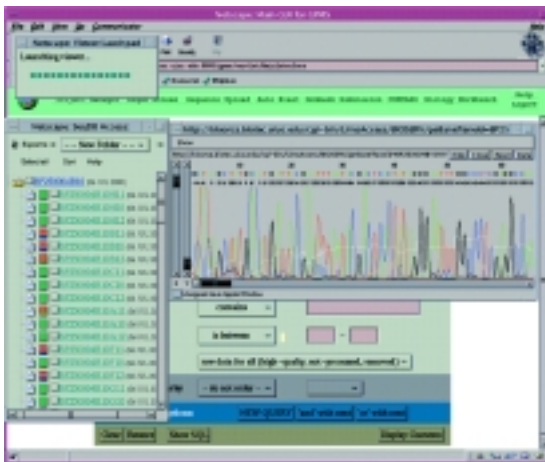
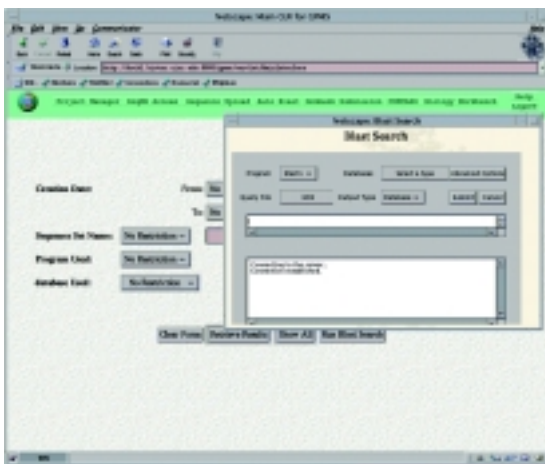


Figure 3. AutoBlast module in the GPMS system. From this module, users can start batch processes of BLAST searches and retrieve results.



results through the built-in user-friendly interface of the program. The Redundancy Check program allows the user to perform redundancy checks on uploaded sequences to evaluate the frequency of identical clones within a project. The Genbank Submission program lets the user submit his/her data to Genbank conveniently. The user can also link to the Biology Workbench and access another suite of software tools for sequence analysis.

GPMS currently hosts and serves several large sequencing projects involving soybean, cattle, bee and bacteria. People can test run the system at the following web site: <http://keck1.biotec.uiuc.edu>. In the future, it will host more campus-wide research projects and serve as an important educational tool in Genomics and Bioinformatics.

—Lei Liu



The Biotechnology Center
103 Observatory
901 S. Mathews Avenue
Urbana, IL 61801
PHONE: (217) 333-1695
FAX: (217) 244-0466
WEB: www.life.uiuc.edu/biotech/

Non-Profit Org.
U.S. Postage
PAID
Permit No. 75
Champaign, IL