

BRC. bioinformatics

Zeng Charts New Territory in Gene Mapping

Dr. Zhao-Bang Zeng, Professor of Statistics and Genetics and member of the Bioinformatics Research Center, has transformed the world of gene mapping by developing statistical methods, along with publicly available software, that use molecular markers to identify locations of multiple genes that contribute to a complex trait, such as size, hardiness, or shape of limbs.

Referred to as quantitative trait locus (QTL) mapping, Zeng's methods also assess how much each QTL affects the trait and how much interaction occurs among QTLs.

QTL mapping has important applications in plant and

animal breeding programs, in human health research, and in the study of the evolution of species.

These methods have helped Trudy Mackay, Professor of Genetics at NC State University, to map genes that are responsible for longevity in *Drosophila*.

Agricultural scientists, such as Charles Stuber, Professor Emeritus of Genetics at NC State, have used QTL mapping to locate genes responsible for the vigor of U.S. corn hybrids. Hybrid vigor is a major contributor to the high yields of this crop in the U.S.

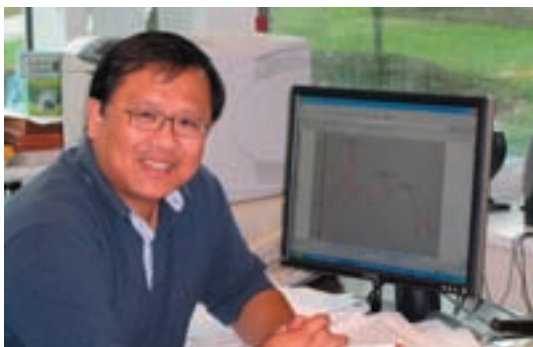
Zeng and researchers at Texas A&M University are now writing an NSF grant proposal to combine QTL mapping with gene expression data from microarray experiments to define genetic networks and determine how genes function together to yield hybrid vigor.

"QTL mapping is the first step in finding specific genes associated with a complex trait. When it is combined with microarray analyses and gene knockout technology, we are no longer mapping genes, we are identifying specific gene function and genetic effects networks," Zeng said.

Continued on page 4

"QTL mapping is the first step in finding specific genes associated with a complex trait. When it is combined with microarray analyses and gene knockout technology, we are no longer mapping genes, we are identifying specific gene function and genetic effects networks."

Zhao-Bang Zeng,
Professor of Statistics
and Genetics



From the Director

The BRC's fifth year has started well, with talented new students, a successful fall retreat, external grant and contract funding at more than \$4 million annually, and authorization to search for a new faculty member.

We were pleased to learn that Provost James Oblinger will assume the position of chancellor on January 1, 2005. As Dean of the College of Agriculture and Life Sciences, Dr. Oblinger has supported the BRC since its inception and has served on our administrative committee. As provost he provided funding to the College of Physical & Mathematical Sciences through the Compact Planning Process for the new faculty position we seek to fill.

Dr. Oblinger reaffirmed the university's commitment to genomic sciences and bioinformatics in his remarks as chancellor-elect on October 8, 2004, and we look forward to working with him to enhance our programs. We wish him

and his wife, Dr. Diana Oblinger, well in their new roles.

Our colleague Zhao-Bang Zeng recently returned from a productive sabbatical at the University of Edinburgh. Some of his groundbreaking work in gene mapping is described in this issue, and future issues will cover the work being done by his large group of students.

In this issue you will also meet three members of our administrative staff, Debra Hibbard, Alex Rogers, and Jb Briseno. The harmonious operation of the BRC is a direct result of their hard work and enthusiasm.

We hope to see you at this year's Cockerham Lecture on November 18, when Dr. James Crow will speak on mutation rates in the human genome. Jim remembers when Clark Cockerham was a graduate student at Iowa State, and he has been a friend of the NC State program since Clark joined the faculty in 1951.

Bruce Weir

"We are pleased to learn that Provost James Oblinger will assume the position of chancellor on January 1. He has long supported genomic sciences and bioinformatics, and we look forward to working with him to enhance our programs."

Sun Shines on Second Fall Retreat

A weekend of perfect weather, river kayaking, and hiking complemented two mornings of research talks by graduate students and faculty at the BRC's second annual fall retreat, held again at the Lakeview Lodge near Bryson City, NC, on October 1–3.

The retreat gives first- and second-year bioinformatics students the chance to learn about research being conducted by faculty and other students and consider the direction their own graduate studies might take.

Participants also enjoy interacting in an informal setting. "The retreat is a really good opportunity for us to interact with first-year students, who mostly involved in classes," said fourth-year student Errol Strain.

"I was able to talk to some of the faculty and get some

ideas on who I might be interested in as an advisor," said first-year student Sam Dickson.

Population genetics, microarray analysis, and quantitative trait loci mapping were common threads in several of the 12 research presentations. Faculty members provided overviews of their work and answered student's question on their research. Students also had a chance to describe the current stages and future directions of their doctoral research.

Kayaking on the Nantahala (with a few bracing spills into the 50-degree water), hiking in the Smoky Mountains, and the now-famous Saturday night barbecue were integral parts of the weekend retreat.

The BRC thanks Spencer Muse for organizing the event once again and contributing so much to its success.



Wen-Ping Hsieh presented her doctoral research at the BRC's 2004 fall retreat.

Langhurst Wins National Gertrude Cox Award

Trena Langhurst, the BRC's trusty work-study student and office assistant since 2002, is one of three winners nationwide of the 2004 Gertrude Cox Scholarship for Women in Graduate Statistics Programs.

The scholarship, named after the first head of the Statistics Department at NC State, is sponsored by the American Statistical Association and the Caucus for Women in Statistics and is offered to exceptional women scholars entering or continuing graduate studies in statistics.

A 2004 Phi Beta Kappa graduate of NC State with a double major in statistics and applied mathematics, Langhurst was named the Outstanding Senior Scholar

and Valedictorian of her graduating class in statistics. She is also a dedicated volunteer, initiating campus-wide collection campaigns for the Marine Corps' Toys for Tots program and Food Bank of North Carolina, among numerous other charitable activities.

Now a doctoral student at NC State, Langhurst is planning a career in social statistics and public service. Both interests stem from her desire to help others overcome problems of domestic violence. "I became interested in how the statistics of domestic violence originate, and began focusing on mathematics, probability, and statistics," she said.

Recent Publications

- Barbash D, Awadalla P, and Tarone AM. 2004. Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus. *PLoS Biology* 2(6):839–48.
- Gibson G, Riley R, Harshman L, Kopp A, Vacha S, Nuzhdin S, and Wayne M. 2004. Extensive sex-specific non-additivity of gene expression in *Drosophila melanogaster*. *Genetics* 167: 1791–99.
- Gibson G and Dworkin IM. 2004. Uncovering cryptic genetic variation. *Nature Reviews Genetics* 5:681–691.
- Haydon DT, Bastos AD, and Awadalla P. 2004. Low linkage disequilibrium indicative of recombination in foot-and-mouth disease virus gene sequence alignments. *J Gen Virol* 85(Pt 5):1095–100.
- Leipzig J, Pevzner P, and Heber S. 2004. The alternative splicing gallery (ASG): bridging the gap between genome and transcriptome. *Nucleic Acids Research* 32(113):3977–3983.
- Maiste PJ, and Weir BS. 2004. Optimal testing strategies for large, sparse multinomial models. *Computational Statistics and Data Analysis* 46:605–620.
- Palsson A, and Gibson G. 2004. Association between nucleotide variation in *Egfr* and wing shape in *Drosophila melanogaster*. *Genetics* 167:1187–98.
- Palsson A, Rouse A, Riley-Berger R, Dworkin I, and Gibson G. 2004. Nucleotide variation in the *Egfr* locus in *Drosophila melanogaster*. *Genetics* 167: 1199–1212.
- Pysz MA, Ward DE, Shockley KR, Montero CI, Conners SB, Johnson MR, and Kelly RM. 2004. Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium *Thermotoga maritima*. *Extremophiles* 8(3):209–17.
- Pysz MA, Conners SB, Montero CI, Shockley KR, Johnson MR, Ward DE, and Kelly RM. 2004. Transcriptional analysis of biofilm formation processes in the anaerobic, hyperthermophilic bacterium *Thermotoga maritima*. *Appl. Environ. Microbiology* 70(10):6098–112.
- Seo T-K, Kishino H, and Thorne JL. 2004. Estimating absolute rates of synonymous and nonsynonymous nucleotide substitutions in order to characterize natural selection and date species divergences. *Molecular Biology and Evolution* 21(7):1201–1213.
- Thorne JL. 2004. Discussant (pages 584–585) on "A Bayesian approach to DNA sequence segmentation," a paper by Richard J. Boys and Daniel A. Henderson. *Biometrics* 60: 573–588
- Weir BS. 2004. Matching and partially-matching DNA profiles. *Journal of Forensic Science* 49(5): 1009–1014.
- Weir B.S., A. Bagdonavicius, B. Blair, C. Eckhoff, C. Pearman, P. Stringer, J. Sutton, J. West and D. Wynen. 2004. Allele frequency data for Profiler Plus loci in Australia. *Journal of Forensic Sciences* 49:112–123.
- Weir, B.S., W.G. Hill and L.R. Cardon. 2004. Allelic association patterns for a dense SNP map. *Genetic Epidemiology* 24:1–9.
- Zhou J, and Gibson G. 2004. Cross-species comparison of genome-wide expression patterns. *Genome Biology* 5(7):232.

Grad Student Probes Hyperthermophile Metabolism

"I've been able to participate in data analysis and in determining relevant biological responses in the data ... There is always a paper in progress, a new experiment in the works, or a new data set that needs a critical eye to interpret."

Shannon Conners,
Bioinformatics
Graduate Student

Shannon Conners, a PhD candidate in bioinformatics and member of Dr. Robert Kelly's research lab, is helping decipher the metabolic mechanisms of hyperthermophiles, a group of microorganisms that thrive at temperatures above 80 C. The organisms hold promise as sources of high-temperature industrial biocatalysts. They might also offer clues about how organisms survive and adapt in extreme conditions, such as might exist on other planets.

Kelly's lab has created full genome cDNA arrays for three hyperthermophiles. Researchers in the lab use differential gene expression data from microarray experiments and functional data from public databases to investigate the physiology and gene function of these organisms. They have tested the microbial responses to carbon source, heat shock, and biofilm growth.

Conners both analyzes and interprets the experimental data. "I am using data sets generated from quantitation of cDNA microarray experiments, with anywhere from 62,000 to 360,000 spot observations," she said.

She analyzes the data using two-stage mixed models as introduced by Russ Wolfinger and Tzu-Ming Chu at SAS Institute and combines the results with func-

tional genomic information from public databases. Then she applies various data filtering and presentation techniques to uncover the biological stories in the data and co-authors the research papers.

Her thesis research combines microarray and sequence data to predict sugar specificities and plausible functions for several large families of sugar-responsive transporters and regulators in the bacterium *T. maritima*. Sequence comparisons with other microbial homologs suggest that several sets of transport proteins have been duplicated in the *T. maritima* lineage and have diverged to fulfill transport capabilities for different sugars.

A biochemistry major at Mount Holyoke College, Conners did two years of graduate work in human genetics at Duke before transferring to NCSU's bioinformatics program in 2000.

"Working in the Kelly group has been a good fit for me because of my biology background," Conners said. "I've been able to participate in data analysis, in identifying relevant biological responses in the data, and in advising other students about interesting research targets. There is always a paper in progress, a new experiment in the works, or a new data set that needs a critical eye to interpret."

The Staff of Life at the BRC

The BRC's continued success reflects the efforts of many people, and among the hardest working are our administrative staff: Debra Hibbard, Julibeth Briseno, and Alexandra Rogers.

Debra is executive assistant, with administrative responsibility for the BRC and for the Summer Institutes in Statistical Genetics. "There's a sense of teamwork here that makes it a positive place to work, and the students enhance that atmosphere," she said. Debra earned her B.S. in Communications at NC State, while working first at the statistics department and then moving to the BRC in 2000. Her friendly manner and "can-do" attitude contribute enormously to the smooth functioning of the BRC and the Summer Institutes. (Trena Langhurst is Debra's right hand in running the Summer Institutes in Statistical Genetics; see her story on page 2.)

Juliebeth (Jb) Briseno is program assistant for the genomic sciences graduate program. Jb manages the flood of applications to the bioinformatics and

functional genomics programs each year and tracks the progress of more than 80 graduate students in the two degree programs. "I have a fun job, because I work with the students," she said. Jb also greets visitors to the BRC, answers the main phone line, and keeps people on their toes with her lively sense of humor.

Alexandra (Alex) Rogers, the BRC's accountant, monitors the financial status of a growing portfolio of grants and contracts and ensures that our faculty, staff, students, and bills get paid on time. As the number of students and employees who are funded by fellowships, industrial internships, and other grants increases, so does her workload. Meanwhile, Alex is pursuing a double major in accounting and business administration at North Carolina Wesleyan College.

The BRC is very grateful to Debra, Jb, and Alex for their uncompromising professionalism, dedication, and good humor.

"I have a fun job, because I work with the students."

Jb Briseno,
Program Assistant,
Genomic Sciences Graduate
Programs

Dr. James Crow will speak on Thursday, November 18 at 4 p.m in Witherspoon Auditorium. A towering figure in the field of population genetics, he is also recognized as an outstanding educator and public servant.

Geneticist James Crow Gives 2004 Cockerham Lecture

Dr. James Crow, Professor Emeritus of Genetics at the University of Wisconsin, Madison, will give the C. Clark Cockerham Lecture on Thursday, November 18 at 4 pm in Witherspoon Auditorium, preceded by a reception at 3:30.

Dr. Crow is a towering figure in population genetics. His books *Genetics Notes* and *Introduction to Population Genetics*, written with Motoo Kimura, defined the field. Jim's

scholarship has been accompanied by outstanding public service, including his work on assessing the genetic effects of radiation and the forensic uses of genetic profiles.

The Cockerham lecture, jointly sponsored by the genetics and statistics departments and the BRC, honors C. Clark Cockerham, a Distinguished University Professor who taught in the statistics department from 1953 to 1990.

Zeng Charts New Territory

Continued from page 1

In wood products research, Zeng and Ron Sederoff, head of NC State's Forest Biotechnology Group, are collaborating on identifying the regulatory genes and genetic pathways that regulate fiber density in eucalyptus, using QTL mapping and whole-genome microarray gene expression data.

Zeng, whose research has been supported by an NIH program project grant since 1990, is noted for two groundbreaking advances in statistical methods of gene mapping. His "composite interval mapping" method, published in 1994, was a significant advance over previous mapping methods and quickly became the standard in the field. Still, the method could evaluate only a single QTL at a time.

In 1999, Zeng and his student Chen-Hung Kao published a "multiple interval mapping" technique that evaluates multiple loci simultaneously, including the

likely magnitude of their effects and the extent of interaction among them.

QTL Cartographer, a free software package first released in 1995, has made the mapping methods accessible to other genetic researchers. Developed by Zeng and his colleagues Christopher J. Basten and Shengchu Wang at the BRC under grants from the USDA, the suite of programs is continuously updated to include advances in statistical gene mapping methods. A Windows version with built-in graphics capabilities and user manual has been available since 2001; a re-designed Windows user manual was released in 2004 (see <http://statgen.ncsu.edu/qtcart/>).

"QTL Cartographer is popular, especially with the new Windows interface. We put a lot of effort into making it easy to use. These methods have made a difference to the research community. That's a major achievement that my colleagues and I are proud of," said Zeng.

"QTL Cartographer has made a difference to the research community. That's a major achievement that my colleagues and I are proud of."

Zhao-Bang Zeng,
Professor of Statistics and
Genetics