VBI e_Connections

feature technology in the news interview

Vol. 2 No. 1 Spring 2007 Newsletter of the Virginia Bioinformatics Institute at Virginia Tech

Why care about prokaryotes?

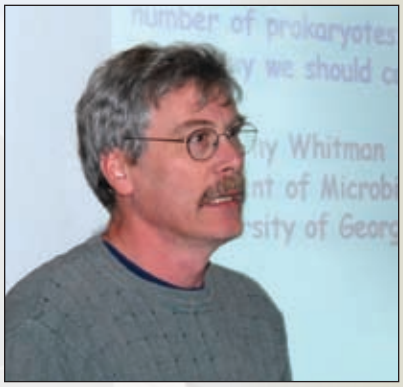
On February 9, Professor William "Barny" Whitman of the University of Georgia Department of Microbiology presented a seminar entitled, "The number and diversity of bacteria on earth and why we care" at the Virginia Bioinformatics Institute Conference Center. In the seminar, Dr. Whitman discussed the biological impact and contributions of prokaryotes, the dominant life form on earth whose evolution established the central plan for the living cell. The presentation was part of the Molecular Cell Biology and Biotechnology (MCBB) seminar series.

In This Issue

As a Ph.D. student at the University of Texas at Austin, Dr. Whitman was one of the first scientists to use a bacterial organism to probe the structure and function of the enzyme ribulose-1, 5-bisphosphate carboxylase (RuBisCO), a catalyst for the chemical reaction by which inorganic carbon enters the biosphere. Dr. Whitman showed that a lysine residue in the enzyme played a critical role in catalysis. Since then, Dr. Whitman has continued his interest in the physiology of environmentally important microbes, and he has studied the relationships of these microorganisms with one another and their surroundings, as well as their metabolic pathways, genetics, phylogeny and systematics. In one study, he and his collaborators set the ambitious goal of counting how many prokaryotes exist today. In May 2000, Dr. Whitman was presented with the 20th Bergey Award, an honor given to an individual for outstanding contributions to prokaryotic taxonomy.

At his seminar Dr. Whitman remarked: "Ancient earth was inhabited by a large population of prokaryotes. The major metabolic processes of today's cells originated in prokaryotes, and animals and plants can trace their history to these organisms. Therefore, it's important to get a handle on their numerical contribution to the biosphere."

He continued: "We have been able to estimate that prokaryotes represent about 3.5 to 5.5 x 10^{17} grams of carbon on the Earth as we know it today, or about the same as the biomass of plants. In terms of individuals, their numbers are about 4 to 6 x 10^{30}. Carbon tracing studies have shown that 3.5 billion years ago, the level of organic carbon in the biosphere was more or less the same. Since the eukaryotes evolved during the last two billion years, it's apparent that the contribution of prokaryotes to the biomass of the early earth was largely prokaryotic."



Dr. William Whitman

The presence of an ancient and large population of prokaryotes has important implications, Dr. Whitman said. All the modern principles of the cell, major biogeochemical cycles and most major lineages of life evolved in a prokaryotic world. Additionally, prokaryotes play a major role in the chemistry of the biosphere.

Dr. Whitman concluded: "Rare genetic events are common in very large populations, and for prokaryotes they have been estimated at 4 x 10^{-7} mutations per gene per DNA replication — thus the potential for diversification is large. It's amazing that genes are conserved at all with numbers like these. The tremendous diversity of prokaryotes has a huge impact on the biosphere of our planet and life as we know it today."

VBI e_Connections

VBI e_Connections is a quarterly publication of the Virginia Bioinformatics Institute produced by the Public Relations team. The newsletter includes feature articles, technology updates as well as interviews that may be of interest to VBI's audiences. Contributions are welcomed.

Please direct submissions to newsletter-editor@vbi.vt.edu
Newsletter team: Susan Bland: Editor; Barry Whyte: Editor; June Mullins: Graphic design.

For further information, please contact:
Barry Whyte at Tel: 540-231-1767, email: whyte@vbi.vt.edu
Website: www.vbi.vt.edu

Roche GS-FLX™ arrives at VBI

The close of 2006 was a busy time for VBI's Core Laboratory Facility (CLF), ushering in exciting, cutting-edge opportunities for the upcoming year and beyond.

In early December, a Roche GS-FLX™ was delivered to the Institute and representatives from Roche and VBI began the installation and training processes. The machine is a genome sequencing system that uses 454 Life Sciences™ sequencing technology, allowing researchers to go from genome to sequence in record time.

The entire installation process took a total of two days, beginning with the task of unpacking all of the individual parts to make sure everything that was needed was available. Two days later, the complete system was up and running, occupying four of the Institute's laboratories. Four rooms are needed to accommodate the technology because there are four separate steps included in the sample preparation process. Each step must occur in a separate room to avoid any issues of contamination. A typical run is completed seven hours after the preparatory steps are finished. Use of the instrument can result in the sequencing of greater than 100 million base pairs per run, and read lengths of 238 base pairs on average. This is possible through the use of a specially designed microplate with wells that only accept a single bead of DNA. The samples are then read by a special light detection device and the signals produced are then quickly converted into a readout of the DNA. This process translates into a cost-effective, high performance solution with many applications

Before the CLF could begin processing samples, however, training was needed and trial runs had to be conducted to ensure the whole system was operating correctly. Clive Evans, manager of VBI's CLF, received five days of training on GS-FLX protocols and how the machine is actually run. He also received an additional day of training on the software package that comes with the system.

"The training was very intensive," Evans explained. "This is a state-of-the-art system and the numerous steps needed to complete a sequence present a challenge to the learning process. Our Roche training representative did an excellent job preparing me to operate the FLX effectively."

Janna Lanza, the Roche representative who led the training efforts at VBI, agreed, adding, "It was a very productive training week. I am confident that VBI will be successful with their new GS-FLX platform."

Pilot runs

A sequencing test run was conducted with the system using a sample of control beads provided by Roche. These samples are designed to be demanding to sequence, serving as an ideal example for a pilot run to test the system. According to Evans, this initial sequencing effort went smoothly. Additional runs were also conducted, including a preliminary run of DNA from *Brucella abortus* strain S19, and an *Escherichia coli* strain provided by Roche. Other preliminary runs were completed that involved processing whole genomes of *B. abortus* S19 and *Helicobacter pylori*. All of these test runs were conducted by Evans after completing training, and the results were highly satisfactory. Additional practice runs have been scheduled for the remainder of this quarter.

"The CLF's goal is to be a one-stop full-service shop for genomic and proteomic services," said Evans. "We pride ourselves in offering the most effective and cutting-edge services, and the addition of the Roche GS-FLX is certainly no exception."

VBI's CLF is a multi-user resource providing various high-throughput technologies and other state-of-the-art technology-related services. The CLF currently provides analysis platforms for DNA sequencing and genotyping, gene expression analysis, and proteomics. Acquisition of the Roche GS-FLX was made possible through Virginia's Commonwealth Research Initiative, which helps universities build their research capacities and stimulate economic development.

Possible future applications

- Shotgun sequencing of genomes – sequencing an entire genome by breaking it into small DNA fragments that can be individually sequenced
- Paired-end sequencing – provides ways to isolate the two ends of a large target nucleic acid into a single small DNA construct for rapid cloning, sequencing, or amplification
- Transcriptome sequencing – analyzing the full complement of activated genes, mRNAs, or transcripts in a particular tissue at a specific time

VBI Buys 454 GS-FLX Sequencer; Will Enable Institute to 'Tackle Whole Genomes'

NEW YORK (GenomeWeb News) — The Virginia Bioinformatics Institute at Virginia Tech has installed a 454 GS-FLX sequencer at its core lab, the Institute said today.

VBI said installing the 454 sequencer will enable its researchers to sequence 100 megabases in 7 hours, and will be able to achieve lengths of as many as 200 base pairs.

Core lab manager Clive Evans said the GS-FLX will help the institute offer the ability to "tackle whole genomes."

VBI, located in Blacksburg, Va., focuses on what it calls the "'disease triangle' of host-pathogen-environment interactions in plants, humans, and other animals."

The core lab provides "various high-throughput technologies" and analysis platforms for DNA sequencing and genotyping, gene expression analysis, and proteomics, VBI said.

Otto Folkerts, associate director of technology development at VBI, said the institute also is excited about new applications that will be coming along for the sequencer, including "sequencing for transcriptomes, clinical samples, paired-end amplicon sequencing and other cutting edge uses" for the sequencer and its related services. Folkerts said these upcoming applications "dovetail well with our current platforms."
Source: GenomeWeb News, January 12, 2007

Life Sciences Organizations Drive Innovation with GraphLogic Technology

GraphLogic, Inc. announced today that 454 Life Sciences, Sequenom, The Virginia Bioinformatics Institute, and Clinical Data are the most recent life sciences companies to license GraphLogic technology to streamline their scientific operations.

GraphLogic's technology - a codeless application builder - provides the benefits of a custom laboratory operations solution without the prohibitive cost and lengthy development time that plague custom systems created with traditional software technology.

The Virginia Bioinformatics Institute has installed a custom microarray application developed by GraphLogic. This web-based system, currently in alpha release, will bring significant advantages of data sharing and data mining to microarray studies.

Calendar of VBI Sponsored Events

- April 19-20 VBI 2nd Annual Research Symposium, Mountain Lake Hotel, Pembroke, VA
- May 3-5 Atlantic Coast Conference on Mathematics in the Life and Biological Sciences, organized by Reinhard Laubenbacher and John Burns, VBI Conference Center
- June 6-8 VBI Annual Faculty Retreat, Berry Hill Plantation, South Boston, VA

See calendar of events at: www.vbi.vt.edu

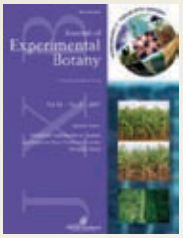
VBI Scientific Publications

The role of neutral lipid nanospheres in *Plasmodium falciparum* haem crystallization

Pisciotta JM, Coppens I, Tripathi AK, Scholl PF, Shuman J, Bajad S, Shulaev V, Sullivan DJ Jr.

Biochemical Journal 2007: **402**(1):197-204.

Researchers at the Virginia Bioinformatics Institute and the John Hopkins Bloomberg School of Public Health have published a study on the role of neutral lipid nanospheres in *Plasmodium falciparum* haem crystallization. When the malarial parasite resides in the red blood cell, it constructs an intracellular haem crystal, called haemozoin, within an acidic digestive vacuole where haemoglobin is degraded. Haem crystallization is the target of the widely used antimalarial quinoline drugs. However, the events that lead to the initiation of the molecular mechanism by which haem crystallization occurs — whether by proteins, polar membrane lipids or by neutral lipids — has not been fully substantiated. In the study, the scientists provide ultrastructural evidence that neutral lipid nanospheres present within the digestive vacuole are the site of hemozoin formation. This may have important ramifications for the design of antimalarial drugs.



Identification of early salt stress response genes in tomato root by suppression subtractive hybridization and microarray analysis

Ouyang B, Yang T, Li H, Zhang L, Zhang Y, Zhang J, Fei Z, Ye Z.

J. Exp. Botany 2007: available on-line at http://jxb.oxfordjournals.org/cgi/reprint/erl258v1 Advance Access published on January 8, 2007.

Work performed at research institutions in China (Huazhong Agricultural University and the National Engineering Research Center for Beijing Biochip Technology) as well as the Virginia Bioinformatics Institute is helping researchers devise ways to tackle high salinity, one of the most serious threats to crop production worldwide. To understand the molecular basis of plant responses to salt stress better, suppression subtractive hybridization and microarray approaches were combined to identify the potential importance of novel genes involved in the early stage of tomato responses to severe salt stress. A total of 201 non-redundant genes that were differentially expressed in early severe salt stress were identified from microarray analysis; most of these genes have not previously been reported to be associated with salt stress. The diversity of the putative functions of these genes indicated that salt stress resulted in a complex response in tomato plants. The use of microarrays will allow researchers to dissect the interplay of these genes in further experiments.

Interview with Athel Cornish-Bowden

Athel Cornish-Bowden is Directeur de Recherche at the Centre National de la Recherche Scientifique, Marseilles, France. Dr. Cornish-Bowden obtained his MA, DPhil and DSc from Oxford University. He is the author of *Fundamentals of Enzyme Kinetics* and *The Pursuit of Perfection*

"I think there was a widespread belief in the 1990s that once the human and other genomes became known a broad understanding of the organization of living organisms would just emerge from the mountain of data. This didn't happen, of course."

Can you briefly describe your career to date? What are your main research interests? I obtained my doctorate at Oxford with Jeremy Knowles, ostensibly in organic chemistry but in practice in enzyme chemistry. I then spent three post-doctoral years with Daniel Koshland at Berkeley, where I developed an interest in the quaternary structure of proteins and its relevance to enzyme regulation. This led naturally to an interest in the principal mechanisms of enzyme regulation, cooperativity and allosteric inhibition. Subsequently I spent 16 years at Birmingham, where I worked on enzyme kinetics and regulation, especially in relation to mammalian hexokinases. Since moving to Marseilles 20 years ago I have become interested in the kinetics of multi-enzyme systems.

Systems biology has been described by some as the United Nations of the biological sciences, which hints at its transdisciplinary nature. What is your working definition for this type of biology? In an ideal world systems biology would be an approach to biochemistry built on a recognition that to understand systems one must study them as systems and not as collections of components. In practice, systems biology seems to differ from traditional biochemistry and molecular biology chiefly in relation to the much larger body of facts that it tries to handle, but true systemic attitudes are not usually to the forefront. Whatever systems biology is, it is a development of biochemistry, not of, say, zoology, botany, taxonomy, etc., or even microbiology.

You have been quoted as saying that "systems biology in current practice is not easy to distinguish from old-style reductionist biochemistry applied on an ever-larger scale." Is systems biology a by-product of our inability to make sense of large data sets? I think there was a widespread belief in the 1990s that once the human and other genomes became known a broad understanding of the organization of living organisms would just emerge from the mountain of data. This didn't happen, of course, and people committed to a systemic view of systems never expected it to happen, so the failure of such an understanding just to "emerge" was certainly one of the driving forces for the explosive growth of systems biology. The term appeared in the scientific literature in 1998, but was used in fewer than 20 papers published before 2000, whereas it occurs in about two papers each day in 2007.

Could you give an example of the limitations of a genome sequencing approach? *Treponema pallidum*, the spirochete responsible for syphilis, now has a completely known genome, but not much more is known about its biochemistry than was known before its genome was sequenced. In the absence of any information about the kinetic and regulatory properties of its enzymes the best we can do is use sophisticated stoichiometric analysis, such as the methods being developed by Stefan Schuster, for example, to deduce from the genome some of its similarities with *Escherichia coli*. That is a step, certainly, but not really what we want to know, which is how it differs from *Escherichia coli*, which does not cause syphilis. Understanding syphilis will require much more than just genome information, and this will include a great deal of kinetic and regulatory information.

Who has had the greatest impact on your career to date? The people who taught me chemistry at Oxford, especially R. J. P. Williams and Jeremy Knowles, had a great impact on my way of thinking about it, as did Daniel Koshland a little later; he in particular encouraged me not to be frightened of building a career on a largely theoretical foundation. In the second half of my career the greatest influence has certainly been Henrik Kacser. He is unfortunately no longer with us, and the other three that I have mentioned have not been primarily interested in systems. Almost everything I have written in the past 20 years has been influenced by discussions with María Luz Cárdenas and Jan-Hendrik Hofmeyr. In the past few years I have been increasingly attracted by the ideas of Robert Rosen, which are also systemic, but at a much more fundamental level than Kacser's.

What advice would you give to a student looking to study biological systems? It would depend on whether the student wanted to make a successful career or wanted to make a real contribution to biological understanding; if the latter, then it would be essential to search for general principles, and to regard accumulating detailed information as a step towards that end and not as an end in itself. Existing methods are quite effective at making people computer-literate; the deficiencies are more on the side of biological literacy. There has always been a widespread feeling that biology is an "easy" subject compared, say, with physics, but this is only true if we just regard it as no more than a vast accumulation of facts. In reality biology needs not just knowledge of a long list of facts, but a broad understanding of how they are related. This is especially important now, at a time when we are witnessing an alarming growth of pseudobiology, a time when creationism, and the supposedly scientific version known as intelligent design, threatens biology teaching throughout the world, no longer just in the United States.

Readers interested in more details are encouraged to consult the following article: "Putting the systems back into systems biology," Perspectives in Biology and Medicine 49.4 (2006) 475-489.