



Bio•IT World

Indispensable Technologies Driving Discovery, Development, and Clinical Trials

**FORWARD THIS ISSUE
TO A COLLEAGUE**



**SUBSCRIBE TO
BIO•IT WORLD MAGAZINE**



Welcome to your digital edition of *Bio•IT World* magazine

We hope the magazine is meeting your expectations. We encourage you to share this issue with your team to stimulate further conversations. [Share this issue.](#)

We are genuinely excited about the prospects for this field—and the magazine—in the coming year. We've cemented our position as the flagship publication of Cambridge Healthtech Institute (CHI), and will continue to play a prominent role in hosting and reporting on the best CHI conferences throughout the year.

We pride ourselves on publishing critical insights and analysis of innovations across the drug discovery pipeline — from molecular modeling of popular drug targets to biomarkers that discriminate cancer responders, and from data handling for next-generation sequencing to new strategies for increasing the speed and efficiency of clinical trials. We will work hard to surpass those stories in the coming year, continuing our pursuit of the most critical tools and strategies that epitomize the world of “predictive biology.”

We hope to continue to engage you with our editorial content, and within our network, and as always, we welcome any and all comments or suggestions —editor@healthtech.com

NAVIGATION KEY

- **Previous/Next Page:** mouse click (or keyboard) on left/right arrows to turn page
- **Contents:** takes you to the Table of Contents page
- **To Zoom:** In the browser version, simply click on the page to zoom in, and again to zoom out; in the enhanced PDF, use the zoom in/zoom out buttons in the top or bottom tool bar.
- **Return to Front Cover:** In the browser version, just click the double arrow button in the Nav bar; in the PDF, click on the Front Cover in the Tool bar.
- **Search:** Click on the search button to perform a full keyword or phrase search. You can search a single issue, or your complete archive.

Go beyond the everyday. **EVERY DAY.**



Symyx Notebook

The Freedom to Experiment.

Whether it's on the trail or in the lab, you want the freedom to take new approaches, routes, and paths to your goals. That's why there's Symyx Notebook. It's the only electronic laboratory notebook that can be deployed across the enterprise in multiple scientific disciplines. With Symyx Notebook, research teams share a single application to document, work, collaborate, and speed the experimentation workflow.

Symyx Notebook streamlines the capture of all experimental information and intellectual ideas. Everyday tasks such as data capture and note taking are optimized and automated. All of which gives you the time and freedom you need to experiment—and get back to doing science.

*To learn more, visit
www.symyx.com/notebook6*

Symposium 2009

Register Now and Save at
www.symyxsymposium.com



© 2008 Symyx is a registered trademark
of Symyx Technologies, Inc. All rights reserved.



Cambridge Healthtech Media Group

www.bio-itworld.com

Bio-IT World

Indispensable Technologies Driving Discovery, Development, and Clinical Trials

MAR. | APR. 2009 • VOL. 8, NO. 2

PAUL BLEICHER
LOOKS BACK AT
PHASE FORWARD **40**

GENOME BIOLOGY
AT MARCO ISLAND **8**

PFIZER'S SEMANTIC
TECHNOLOGIES **34**

SOCIAL NETWORKS
FOR SCIENTISTS **16**

COMPUTING for CANCER

Dana-Farber's
John Quackenbush on
How to Merge and Mine
Patient Data **28**

#1

Clinical Trial Management System

Get Better Results With Oracle

ORACLE®

**oracle.com/goto/healthsciences
or call 1.800.ORACLE.1**

Copyright © 2008, Oracle. All rights reserved. Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

Successfully delivering IT and consulting services to world leaders in R&D, science and engineering



Innovative Informatics and Consultancy for R&D

For over 25 years Tessella has worked with some of the biggest names in the pharmaceutical industry. We help clients address a wide range of business goals, including, improving efficiency and productivity in drug discovery, delivering the business benefits of translational medicine, and faster more effective clinical trials.

What makes our technology and consulting services effective is our understanding of both the science and informatics of pharmaceutical R&D, as well as unique software assets and intellectual property that reduce project risk, cost and time.

With offices in Europe and North America, global companies rely on Tessella for business critical assignments.

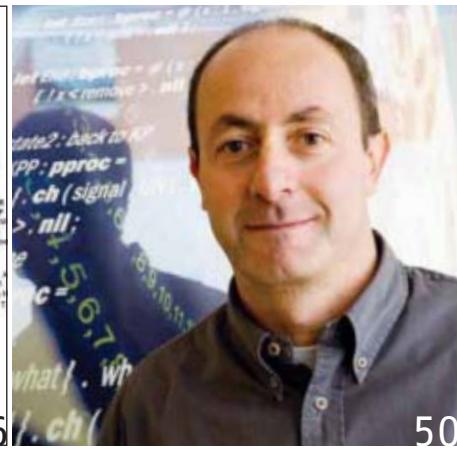
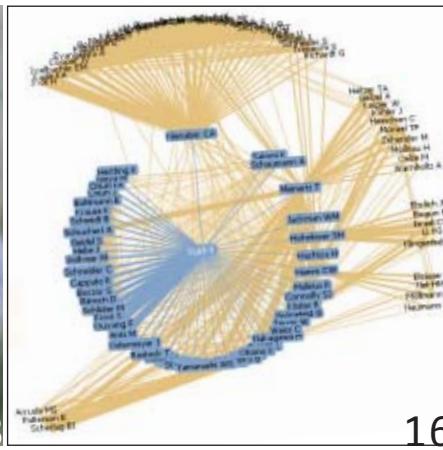
- **Greater efficiency and productivity in drug discovery**
- **Delivering the benefits of translational medicine**
- **Faster, more effective clinical trials**

**Visit us on booth # 203
Bio-IT World Expo
April 27-29th 2009
Boston, MA**

www.tessella.com
info@tessella.com

 **Tessella**
Technology & Consulting

Contents [03–04·09]



Cover Story

Integrating Clinical and Genomic Data

With assists from Oracle and InforSense, Dana-Farber's John Quackenbush finds ways to mine vast patient data collections to better understand human cancer. **28**

Up Front

- 8 "Breathtaking" Biology Shimmers on Marco Island**
At AGBT, new technology heralds exciting new biology.
- 9 Sage Launched at MMTC**
Merck execs tout new venture as key ingredient for disease biology.
- 12 Are Patent Offices Ready for Personalized Medicine?**
New science presents unique challenges to the patent system.
- 14 The State of Kits**
Inside the Box We may be ready for an open source, next generation toolkit. **BY BRIAN OSBORNE**
- 15 Expanding Molecular Diagnostics**
Insights Outlook Molecular diagnostics has exciting opportunities.
BY LUCY SANNES
- 9 Briefs**

Computational Biology

- 16 Keeping Science Connected**
The rise of scientific social networking is paying dividends for users and entrepreneurs alike.

Computational Development

- 19 Chasing Cellular Systems Biology**
Cellumen uses high content screening to flag tox-causing compounds.
- 20 Agile IT Strategy in Pharmaceutical R&D**
Clearly communicated aims should be the goal.

IT/Workflow

- 34 Ted Slater's Semantic Technologies**
The semantic web doesn't exist, but Pfizer's Slater believes semantic technologies are paving the way.
- 36 Platforms for Personalized Medicine**
Integrating clinical information and R&D data will ensure success.

Clinical Research

- 40 A Web MD: Paul Bleicher Reflects on Phase Forward**
Bleicher talks about Phase Forward's acquisitions, milestones, and what's next.

- 42 MacGarvey Establishes Quanticate's U.S. Operation**
The Cambridge office expands reach of clinical trial services and products in North America.

In Every Issue

- 5 Collaboration and the Long Tail of Disease**
First Base Jay Tenenbaum urges pharma to share resources.
BY KEVIN DAVIES
- 50 CoSBi Models**
The Russell Transcript Corrado Priami, president of CoSBi, hopes to infer principles from living systems to improve computer science.
BY JOHN RUSSELL
- 6 Company Index**
- 6 Advertiser Index**
- 46 Educational Opportunities**
- 48 New Products**

Bio-IT World

SPECIAL ADVERTISING SECTION

Next-Generation Sequencing and Diagnostics

Begins on page 23

First Base



Collaboration and the Long Tail of Disease

KEVIN DAVIES

Ten years ago, Jay “Marty” Tenenbaum, a highly successful Internet entrepreneur, was diagnosed with metastatic melanoma and given 12 months to live. He researched various experimental drug treatments, and credits a failed cancer vaccine, among other drugs, for saving his life. Through the company he founded, CollabRx, Tenenbaum aims to leverage the extraordinary untapped expertise and resources across the industry to empower individual patient health care through personalized research.

In a powerful opening keynote at CHI’s Molecular Medicine Tri-Conference (MMTC) in San Francisco in February, Tenenbaum, founder and chairman of CollabRx, urged members of the life sciences community to share their resources to empower personalized research and help satisfy the unmet medical needs of the “long tail” of disease. “As a patient... I want to tap all of the world’s knowledge and all of the world’s resources into curing my disease,” Tenenbaum said, his voice ringing with commitment. “I was there in the very early days of e-commerce... the vision is absolutely clear. For those of you who get it, there is the potential to become the Microsoft’s, the Google’s, the eBay’s, the Amazon’s of this industry... For those of you who don’t get with it, unfortunately you’re going to wind up like Waldenbooks or Egghead Software or Encyclopedia Britannica.”

Tenenbaum argued that individual patients, such as those suffering severe forms of cancer, cannot wait for typical drug trials that take 15 years and cost more than \$1 billion. But that time can be cut from years to months by slashing clinical trials, replacing group statistics with deep, genomic profiling of the tumor, including whole-genome sequencing, to produce “a very detailed picture of the biology and the pathways driving this person’s disease.”

Moreover, rather than spend years identifying new drug molecules, CollabRx research focuses on the thousands of drugs that have already passed FDA approval. Computational tools can then map those existing drugs, used off label or in cocktails, to find combinations to help individual patients. The success of CollabRx in finding therapies for many patients rep-

resented “a new gold standard... in which every patient gets the benefit of the best available science.”

CollabRx ONE

Tenenbaum announced the launch of a new personalized research service for oncology called CollabRx ONE. Patients can instruct hospitals to send biopsy samples to CLIA-approved labs specified by [CollabRx](#), for detailed genomic analyses of the tumor. Those data will be analyzed by computational and systems biologists and interpreted for the benefit of the patient’s oncologists who may not be versed in molecular genetics.

“Within weeks to a month of starting this process, we can get a drug into a patient. And we can learn within weeks to a month after that, based on either biomarkers or imaging, whether or not that patient is responding to therapy. This is really unbelievable.” Tenenbaum called it “real-time research—there’s no daylight between research and the patient.”

Tenenbaum supports several virtual biotechs including one for the [Melanoma Research Alliance](#), which is helping to classify the subtypes of melanoma. Cancer samples are being distributed to the [Broad Institute](#) and [TGEN](#) for profiling. From there, oncologists are setting up virtual trials networks on targeted sub-populations of patients. The goal is to marry the virtual biotech and the personalized research service, in order to validate the process on 50 to 100 patients.

“If you’re a company or a researcher or an oncologist or a patient even who is involved or concerned with melanoma, you’ve got to be connected to a network like this,” Tenenbaum says.

But to succeed, Tenenbaum needs help. “I need specimens, I need drugs, I need access to screening

libraries, mouse models, and laboratory facilities.” A number of organizations started [Health Commons](#), including [Science Commons](#), the [Public Library of Science](#), and Tenenbaum’s [Commerce Net](#). New collaborators include the Personal Genome Project, Treat 1000, and a brand new entity called Sage, created by Steven Friend and Eric Schadt of [Merck](#) (another big MMTC announcement—see p. 9).

Tenenbaum wants researchers and industry executives to consider the unused assets sitting on shelves that are not being monetized, but yet “might be the key to saving someone’s life.” A personalized research service can provide answers to patients within weeks to months, he concluded. “This is the way patients with serious diseases are going to be treated... We’re doing it today, and it’s only going to get better.”



Jay Tenenbaum

Company Index

454	8	Harvard	18	Public Library of Science	5
Agilent	10	Health Commons	5	Quantitate	42
Applied Biosystems	8, 48	Helicos Biosciences	8	Roche	8
Beijing Genomics Institute	8	H. Lee Moffitt		Sage	5
Bioinformatics Organization	9	Cancer Center	10, 29	Science Commons	5, 9
Broad Institute	5, 8, 9	IBM	9	SciLink	16
Cambridge Healthtech		Illumina	8	Stanford	18, 10
Associates	22	Indiana University	9	Symyx	48
Celera Genomics	32	InforSense	31	Tessella	22
Cellumen	19	Johns Hopkins University	18	TGEN	5
Centre for Computational and		Melanoma Research Alliance	5	The BioTeam	14
Systems Biology	50	Merck	5, 9	The Institute for Genomic	
CollabRx	5	MIT	18	Research	29
Collexis	16	National Human Genome		University of California, Riverside	34
Commerce Net	5	Research Institute	8, 32	University of California,	
Complete Genomics	8, 9	Navigenics	9	San Francisco	18
Dana-Farber Cancer Institute	29	NVIDIA	48	University of California, Davis	9
DeLano Scientific	9	Oracle	30, 36	University of California,	
Elsevier	16	Osmetech	15	San Diego	9
Entelos	50	Oxford Nanopore Technologies	10	University of Maryland	9
Epernicus	18	Pacific Biosciences	8	University of Trento	50
GCG	34	Paradigm Genetics	34	University Washington	10
Gene Network Sciences	50	Penn State	10	Waters	48
Genstruct	50	Pfizer	34	Welcome Trust Sanger Institute	9
Genzyme	18	Phase Forward	40	Yale	9

Advertiser Index

Advertiser	Page #	Advertiser	Page #
Accelrys www.accelrys.com	24-25	Educational Opportunities www.bio-itworld.com	46-47
Adobe Webcast www.bio-itworld.com	21	Eli Lilly Singapore www.lsccd.lilly.com.sg/lsccd/careers	27
Barnett Educational Services www.barnettinternational.com	37	ERT www.ert.com	41
Bio Georgia www.georgiabiosciences.com	11	GeneGo www.genego.com	Cover 3
Bio-IT World 2009 Best Practices Awards Dinner www.bio-itworldexpo.com/awards.asp	49	Google Webcast www.bio-itworld.com	45
Bio International Convention www.convention.bio.org	33	Insight Pharma Reports www.insightpharmareports.com	43
Bio-IT World Conference and Expo 2009 www.Bio-ITWorldExpo.com	38-39	LabVantage www.labvantage.com/suite	7
Biomarker Bridge www.biomarkerbridge.com	17	Oracle www.oracle.com/goto/healthsciences	Cover 2
Bluearc www.bluearc.com/html/infocenter	13	Symyx www.symyx.com/notebook6	Cover 4
CLC Bio www.clcbio.com/genomics	26	Tessella www.tessella.com	3
eCliniqua www.ecliniqua.com	35		

This index is provided as an additional service. The publisher does not assume any liability for errors or omissions.

VOLUME 8, NO. 2

Editorial, Advertising, and Business Offices: 250 First Avenue, Suite 300, Needham, MA 02494; (781) 972-5400

Bio-IT World (ISSN 1538-5728) is published bi-monthly by Cambridge Bio Collaborative, 250 First Avenue, Suite 300, Needham, MA 02494. Bio-IT World is free to qualified life science professionals. Periodicals postage paid at Boston, MA, and at additional post offices. The one-year subscription rate is \$199 in the U.S., \$240 in Canada, and \$320 in all other countries (payable in U.S. funds on a U.S. bank only).

POSTMASTER: Send change of address to Bio-IT World, P.O. Box 3414, Northbrook, IL 60065. Canadian Publications Agreement Number 41318023. CANADIAN POSTMASTER: Please return undeliverables to Station A, P.O. Box 12, Windsor, ON N9A 6J5.

Subscriptions: Address inquiries to Bio-IT World, P.O. Box 3414, Northbrook, IL 60065 (888) 835-7302 or e-mail biw@omedamedia.com.

Reprints: Copyright © 2009 by Bio-IT World All rights reserved. Reproduction of material printed in Bio-IT World is forbidden without written permission. For reprints and/or copyright permission, please contact the YGS group, 1808 Colonial Village Lane, Lancaster, PA; 717-399-1900 X 125 or via email to ashley.zander@theYGSgroup.com.

Bio-IT World®

Indispensable Technologies Driving
Discovery, Development, and Clinical Trials

EDITOR-IN-CHIEF

Kevin Davies (781) 972-1341
kevin_davies@bio-itworld.com

MANAGING EDITOR

Allison Proffitt (781) 972-1345
aproffitt@healthtech.com

ART DIRECTOR

Mark Gabrenya (781) 972-1349
mark_gabrenya@bio-itworld.com

VP BUSINESS DEVELOPMENT

Angela Parsons (781) 972-5467
aparsons@healthtech.com

VP SALES – WESTERN US, CANADA, EUROPE, PACIFIC RIM

Alan El Faye (213) 300-3886
alan_elfaye@bio-itworld.com

REGIONAL SALES MANAGER – NEW ENGLAND,
NORTH EASTERN US, SOUTH EASTERN US,
MIDWEST, INDIA

Kay O. Christopher (860) 693-2991
kchristopher@healthtech.com

SENIOR DIRECTOR OF MARKETING &
OPERATIONS, PUBLICATIONS

Joan A. Chambers (781) 972-5446
jchambers@healthtech.com

PROJECT/MARKETING MANAGER

Lynn Cloonan (781) 972-1352
lcloonan@healthtech.com

ADVERTISING OPERATIONS COORDINATOR

Matt Gross (781) 972-5481
mgross@healthtech.com

PRODUCTION MANAGER

Tom Norton (781) 972-5440
tnorton@healthtech.com

Contributing Editors

Michael Goldman, Karen Hopkin,
Deborah Janssen, John Russell,
Salvatore Salamone, Deborah Borfitz
Ann Neuer, Tracy Smith Schmidt

Advisory Board

Jeffrey Augen, Mark Boguski,
Steve Dickman, Kenneth Getz,
Jim Golden, Andrew Hopkins,
Caroline Kovac, Mark Murcko,
John Reynders, Bernard P. Wess Jr.

Cambridge Healthtech Institute

PRESIDENT
Phillips Kuhl

Contact Information

editor@healthtech.com

250 First Avenue, Suite 300
Needham, MA 02494



A WORLD OF LABORATORY INFORMATION MANAGEMENT
OUT OF THE BOX



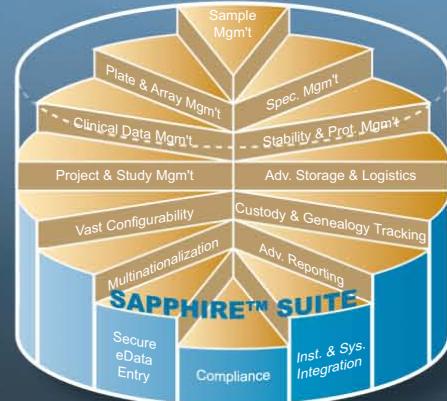
Sapphire™ Suite: The Best Combination of Functionality, Flexibility and Technology

Successful organizations can no longer depend on one-dimensional systems. As your business grows, so must your systems and their functionality. LabVantage redefines the "S" in LIMS with the **Sapphire Suite** -- a zero-footprint and configurable-off-the-shelf Laboratory Information Management Suite. Sapphire takes you beyond the world of traditional LIMS by consolidating additional functionality typically found in multiple laboratory systems into one product platform across the global enterprise. Sapphire is the best combination of functionality, flexibility, & technology in one solution.

Configure your success with the Sapphire™ Suite at:

www.labvantage.com/suite

Keyword: OutOfBoxLIMS



LabVantage is an IBM® Business Partner, who has demonstrated success in delivering thin-client laboratory information management solutions for biosciences and pharmaceutical customers. LabVantage's Sapphire runs on IBM on demand software and hardware infrastructure including Linux® software platform and System x. This proven technology is tailored to address the business and IT needs of research, development, formulation, and quality management laboratories. To find out how you can leverage IBM's on demand technology for success in laboratory information management, visit: www.ibm.com/lifesciences.



AUSTRIA • BELGIUM • CANADA • CHINA • DENMARK • ESTONIA • FINLAND • FRANCE • GERMANY
 HONG KONG • INDIA • JORDAN • LATVIA • LITHUANIA • LUXEMBOURG • MALAYSIA • NETHERLANDS
 NORWAY • OMAN • QATAR • SAUDI ARABIA • SINGAPORE • SOUTH AFRICA • SWEDEN • SWITZERLAND
 TAIWAN • THAILAND • UNITED ARAB EMIRATES • UNITED KINGDOM • UNITED STATES

Up Front News

'Breathtaking' Biology Shimmers on Marco Island

At the 10th anniversary of AGBT, new technology heralds exciting new biology

BY KEVIN DAVIES

MARCO ISLAND, Florida—While the scientific debut of [Complete Genomics](#) captured much of the attention at this year's sold-out AGBT conference*, there were plentiful signs that next-generation sequencing technology is delivering much more than mass human genomes.

Opening the conference, [Broad Institute](#) director **Eric Lander** saluted the "intellectual and commercial ferment" in genome analysis that is not only generating unprecedented volumes of data but also spurring discovery. "It's breathtaking, an entire community, academic and commercial, coming together," raved Lander. Current next-generation sequencing instruments are churning out 2 billion bases (Gb) of sequence per day, numbers that will look puny next year. [Illumina](#), [Applied Biosystems](#), and [Roche/454](#) all presented impressive data showing a doubling of throughput in 2008, with even more dramatic increases on tap.

Lander offered examples of how high-throughput sequencing is becoming a routine tool for tackling molecular biology problems. "For each disease, we're going to have to sequence thousands of patients," he said, to identify the rare causal variants. In February, Lander, John Rinn, and colleagues published a landmark discovery by studying chromatin state-maps to identify long intergenic non-coding RNAs (lincRNAs). Analysis of a telltale pattern of histone modifications—"If you see a K4 and a K36 [modification], it marks a gene"—delineates a staggering 1600 lincRNAs (up from the textbook tally of 12!). There is already evidence suggesting some have functional roles in gene regulation.

Les Biesecker ([NHGRI](#)) described

*Advances in Genome Biology and Technology, Marco Island, Florida; February 4–7, 2009.



SAVING ENDANGERED SPECIES SUCH AS THE TASMANIAN DEVIL IS A NEW TWIST FOR NEXT-GEN SEQUENCING.

ClinSeq, an ambitious study that launched two years ago to generate detailed sequence information on 400 candidate genes in 1000 subjects. "We can't get patients through the clinical center as fast as they want to enroll," said Biesecker. So far, his group has sequenced more than 825 megabases (Mb). In one patient with high cholesterol and a "stupendously high coronary calcium" level, the study revealed a stop mutation in the LDL receptor gene.

Wang Jun ([Beijing Genomics Institute](#), Shenzhen), said his institute is sequencing 20 Gb per day with a fleet of 18 Illumina GA II machines and a data center with 1000 CPUs and 1500 terabytes (TB) storage. Wang is selecting 100 species with little competition and a significant Chinese element, such as the giant panda, under assembly using in-house algorithms based on the Bruijn graph theory.

RUMORS AND PROMISES

According to CSO **Steve Turner**, [Pacific](#)

[Biosciences](#)' debut instrument won't be available until late 2010, but there is healthy progress among the first prototypes. Last November, PacBio sequenced a 107-kb stretch of human DNA (in a bacterial artificial chromosome). The average read length was 446 bases, with some reads exceeding 2000 bases. Accuracy was 99.99% in the non-repeat regions. In January, PacBio sequenced *Escherichia coli* (38x coverage), finding just five discrepancies in the genome.

After a difficult year with job cuts, management reshuffles, and meager sales, [Helicos Biosciences](#)' **Bill Efcavitch** declared, "The rumors of our demise are greatly exaggerated." The HeliScope is producing up to 150 Mb/hour, with average read lengths of 30 bases. In a test sequence of the nematode, Efcavitch said the total error rate was 3.5%, and would move on to attempt a human sequence.

In a much anticipated talk from the CEO of Complete Genomics, **Clifford Reid** announced the release of his company's first assembled human genome (250 Mb mappable reads). Reid said the quality—a discordant rate of 0.34%—was equal to the published African and Asian genomes. He anticipated achieving 200 Gb/run this summer, and triple that by the end of 2009, while still shooting for the \$5000 genome later this year. A service model was the only way to go: "We'll sequence, assemble, generate the variants, and send it back to you. 60,000 processors are going to light up," he joked. (*Editor's Note: Reid keynotes Bio-IT World Expo on April 29, 2009.*)

John Todd (Cambridge University) described progress in type 1 diabetes (T1D). Genome-wide association studies have implicated dozens of candidate genes, but which ones are causal? Deep sequencing of one such gene, *IFIH1*,

(CONTINUED ON PAGE 10)

Briefs

Merck Execs Tout Sage as Key Ingredient for Disease Biology

Landmark announcement at CHI's MMTC signals open-access initiative for drug development

BY KEVIN DAVIES

SAN FRANCISCO—**Merck** executives Stephen Friend and Eric Schadt unveiled their plans for Sage, an open access platform for sharing and disseminating complex data representing disease biology, at CHI's Molecular Medicine Tri-Conference last month.* In a joint presentation, Friend, Merck senior vice president and former oncology chief, and Schadt reviewed the successes and outstanding challenges that prompted them, with Merck's blessing (in the form of money and resources) to entertain a bold new approach to improving the expense, time, and productivity of drug development.

The benefits of analyzing complex biologics networks are very good, said Schadt, but “more expensive than any one company can afford.” The vision of Sage was “to create open access, integrative bionetworks, evolved by contributor scientists, to accelerate the elimination of human disease.” An all-star advisory team includes Nobelist Leland Hartwell, Sir David Lane (A*STAR Singapore), Navigenics co-founder Dietrich Stephan, Merck research chief Peter Kim, Yale’s Rick Lifton, and John Wilbanks ([Science Commons](#)).

Schadt said that Sage would be absolutely dependent on contributing scientists across the globe. “We need massive amounts of information appropriately integrated to build models that are predictive,” said Schadt. Aside from the scale and cost of such research, “scientists across the globe involved in different areas of research need to be actively engaged in accessing these networks and contributing information back.”

The transition from a linear to a network mindset would require the generation of coherent datasets, the development of predictive models to design

novel therapeutic approaches, and the leveraging of social networks and other means to foster a contributor network. “Watching the trends of public data access, we anticipate a transition of disease biology into the precompetitive space,” said Schadt. Friend added that the transition is, “something that we feel in the long run has an opportunity [to succeed].” As for why scientists should include their own data, Friend said, “picture chemistry, picture physics. The people who were originally trying to mix compounds didn’t get very far until they found molecular structures... This is the analogy for what’s going to happen in biology.” New representations of disease allow for data to be shared and layered.

“The hardest part may not be the technology,” Friend concluded. “It’s either going to be ... our institutions ... that have a certain culture about what we do with data. Or it’s going to be the clinicians,” who aren’t used to presenting clinical data using defined standards.

Decade of Discovery

Over the past decade, Friend said Merck has introduced numerous bold technologies that have been successful in limited capacities, including widespread RNA expression profiling in tumors, which led directly to the development of Mammaprint and Oncotype diagnostic tests; whole-genome RNA interference (RNAi) screening to select drugs and patient response. But system and sample heterogeneity made it almost impossible to put the results into context. “It’s like looking at a single frame in Slumdog Millionaire and going, Ah, that’s what that movie was about,” he said.

A third initiative, beginning around 2002, was to merge databases of clinical information and genetic information.

(CONTINUED ON PAGE 10)

FRANKLIN AWARD NOMINEES

The [Bioinformatics Organization](#) has announced the names of the six nominees for the 2009 Benjamin Franklin Award. The winner will be deliver a lecture at the Bio-IT World Conference & Expo, on Tuesday, April 28. The nominees are: Philip E. Bourne (Co-Director, Protein Data Bank, [University of California San Diego](#)); Warren DeLano ([DeLano Scientific](#)); Jonathan Eisen ([University of California, Davis](#)); Don Gilbert ([Indiana University](#)); Heng Li ([Wellcome Trust Sanger Institute](#)); and Steven Salzberg ([University of Maryland](#)).

SEQUENCING COLLABORATION

[Complete Genomics](#) (CGI) is collaborating with the [Broad Institute](#) to conduct complete human genome sequencing pilot studies of two cancers—glioblastoma and melanoma. The Broad Institute will provide CGI with five cancer samples, including one that has been extensively characterized by the scientific community. The other four genomes are matching glioblastoma and melanoma tumor/healthy tissue pairs.

GENOMICS SHIFT

[Navigenics](#) CEO Mari Baker has stepped down to join PlayFirst, an entertainment and video game company, as CEO. Baker will continue to serve on the Navigenics Board of Directors, and in the interim, board members Dana G. Mead, Jr. and Sue Siegel will work with the existing executive team to manage the business.

LIFE SCIENCES PARTNERSHIPS

A survey by [IBM](#) and Silico Research found that although the number of alliances between larger and smaller life sciences companies has increased, the partnerships aren’t producing new medicines. 55% of life sciences CEOs do not plan to partner extensively over the next three years.

*CHI's Molecular Medicine Tri-Conference, San Francisco, February 23–26, 2009.

Up Front News

Marco Island

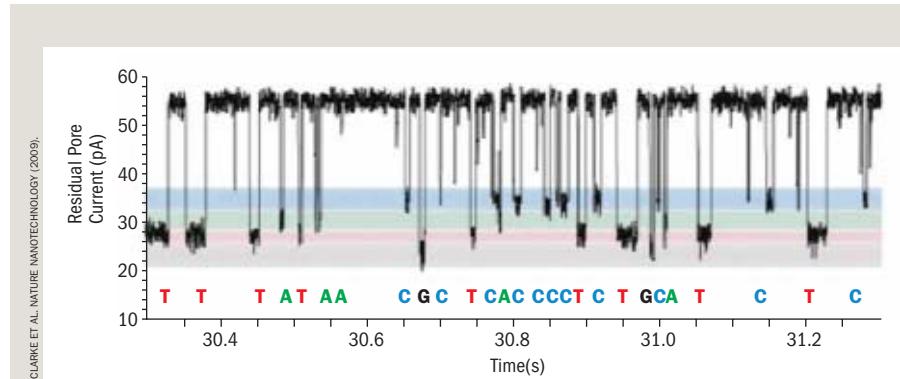
(CONTINUED FROM PAGE 8)

uncovered four mutations, prompting follow-up studies in 10,000 patients. "The results were fantastic," said Todd. *IFIH1* turns out to confer protection when it harbors mutations. "When functional, you're susceptible," he said. More exciting still, the gene encodes an intracellular receptor for the coxsackie virus—a known risk factor in T1D.

Former George Church protégée **Jay Shendure** (Univ. Washington) used Agilent arrays to capture 25 Mb coding (exome) sequence of ten humans before sequencing. Included were two patients with a rare genetic disorder, Freeman-Sheldon syndrome. Shendure demonstrated that exome sequencing could pick up Mendelian mutations—21 genes contained novel variants, but only the mutation in *MYH3* (the known disease gene) was predicted to be damaging.

Final Four

The closing quartet of speakers began with Stanford Nobelist **Andy Fire**, who is using 454 technology to sequence immunoglobulin genes, despite admitting to "an F in immunology as a graduate student." He hopes to use patients' hypermutation status for a rapid prognostic cancer test. **Bruce Budowle** (FBI) said that biowarfare is nothing new, and dates back to the ancient Romans. His group uses SOLiD sequencing to sequence suspected anthrax and Yersinia strains. **Len Pennachio** described a wealth of data (see enhancer.lbl.gov) for identifying enhancer elements and characterizing their expression sites. Closing the conference was **Penn State's Stephan Schuster** on ancient DNA and museomics. His group has sequenced the DNA of extinct and endangered species, notably the Tasmanian Devil, which is severely threatened by an infectious cancer outbreak. Schuster is sequencing Cedric—voted "Tasmanian of the Year" in 2008 for surviving test infections. The goal, said Schuster, is to use SNP information to "direct the breeding program and do pedigree selection in insurance populations." In other words, increase the genetic variation to increase fitness. "We are racing to the finish line." •



Sharp dips in current discriminate the four bases of DNA passing through the nanopore.

Oxford's Opening Statement

Oxford Nanopore Technologies recently published its first proof of principle of its label-free next-generation sequencing technology in *Nature Nanotechnology*. The paper*, by James Clarke and colleagues, shows that an engineered nanopore can discriminate between the four bases of DNA with remarkable specificity, and can even identify methylated C residues ("the fifth base"). Senior author is Hagan Bayley, Oxford University chemistry professor and co-founder of Oxford Nanopore.

Clarke's team genetically engineered the alpha-hemolysin to covalently attach cyclodextrin—a washer that sits in the middle of the pore. By altering variables such as salt concentration, pH and temperature, Clarke could resolve the four bases to an accuracy of about 99.8%. The slight difference in size of the four bases results in a discrete blockage of the pore, which is read as a reproducible dip in the applied current from a baseline of around 60 picoamps to anywhere from 20–40 picoamps.

The authors conclude: "These advances represent the realization of a complete nanopore base detector, which, when combined with a compatible exonuclease DNA processing system, will provide the basis of a complete nanopore sequencer."

That is the next step—to couple an exonuclease to the nanopore so that single bases are cleaved off a template DNA strand and their identity read off instantaneously as they

Editor's note: Oxford Nanopore's James Clarke will be presenting a poster on this work at CHI's Next-Generation Sequencing conference, March 17–19, in San Diego.

*Clarke J. et al. "Continuous base identification for single-molecule nanopore DNA sequencing." *Nat Nanotech* 2009

funnel through the nanopore. The *Nature Nanotechnology* paper also shows that the nanopore can detect bases cleaved off DNA strands in solution by the enzyme.

Sage

(CONTINUED FROM PAGE 9)

Merck forged collaborations with institutes such as the **Moffitt Cancer Center** (see, "**Merck-Moffitt Partnership Breaks Down Silos**," *Bio•IT World*, Aug 2008), which enables Merck researchers to direct patient selection in clinical trials based on molecular signatures. But Friend said that the volume of disease data amounted to "a clinical/genomic Tower of Babel" problem.

More recently, Merck has been riding the success of Schadt's team in Seattle (see, "**Eric Schadt's Integrative Approach to Predictive Biology**," *Bio•IT World*, Oct 2008), which has taken major steps to harness the explosion of data and analyze biological networks to predict the physiological state of the system. "To be competitive in the future and to impact human health, we must become masters of information," Schadt said, displaying a picture of Aria, the all-seeing master computer from the film Eagle Eye. •



Where bioscience
companies have
global reach

No wonder Georgia is hosting the 2009 BIO International Convention. In Georgia, world-class research, preeminent universities, and global accessibility accelerate almost 300 bioscience companies. Georgia is among the top 10 states for research and development, thanks to institutions like Emory University, Georgia Tech, and the University of Georgia. It's also home to the world's top public health agency, the U.S. Centers for Disease Control and Prevention. Your company will enjoy easy access to the world through Hartsfield-Jackson Atlanta International Airport.

For an up-close look at what Georgia has to offer, call 404.962.4040
to schedule your VIP site tour. Visit www.georgiabiosciences.com.



Crossroads of Global Health

Up Front News

[GUEST COMMENTARY]

Are Patent Offices Ready for Personalized Medicine?

New science presents unique challenges to the patent system.

BY R. BRIAN MCCASLIN

Politicians, physicians, health insurers, and drug companies all see personalized medicine as a panacea that simultaneously will improve patient care, reduce health care and development costs, and ensure profitability. While the various constituencies have rallied around the potential of personalized medicine, various patent offices have not updated their procedures to accommodate the technologies driving personalized medicine.

Fundamental tenets of personalized medicine include the elucidation of an individual's risk of developing a particular disease state and an assessment of the person's likelihood of responding to a certain therapy. While the relevant assays can involve measuring a single, well-defined biomarker, they more likely comprise an evaluation of multiple biomarkers whose identities or function may not be known.

A suite of mass spectrometry platforms is typically used to identify and measure markers indicative of a disease state or susceptibility to therapy. Unfortunately for innovators striving to develop new proteomic assays, the current rules and procedures promulgated by patent offices were crafted in view of the old paradigm assays and have proven troublesome for the new technologies.

Prosecuting Proteomic Inventions

For example, current implementation of restriction and unity practice in, respectively, the United States Patent and Trademark Office (USPTO) and the European Patent Office (EPO), makes prosecuting the full scope of proteomic inventions prohibitive.

Consider the fictitious company BIORite, which employed tandem mass spectrometry and pattern recognition

algorithms to identify 200 biomarkers for schizophrenia. BIORite is anxious to commercialize its technology and prepares a patent application with claims directed to a method of assessing schizophrenia in a patient by evaluating a clinical sample for the presence of at least one of biomarkers 1-200, which are identified by molecular weight.

More than likely, the USPTO and EPO will assert that such a claim encompasses 200 inventions which must be pursued in separate (i.e. 200) patent applications. Likewise, if BIORite decides to pursue foreign protection via the Patent Cooperation Treaty (PCT), the PCT examiner will likely consider the claims to encompass 200 inventions and require BIORite to pay \$1,000 per additional invention (i.e. \$199,000) to fully examine the application.

BIORite will also face an uphill battle to secure patent protection for each of the restricted inventions. Seeking to impose the perspective of the former paradigm on the new platform technologies, patent examiners are apt to reject the new claims on a variety of fronts.

For example, in the U.S., examiners may reject the claims (under 35 U.S.C. §112, ¶1) for allegedly lacking "written description." European examiners could cite Article 84 for an alleged lack of clarity. In both instances, the rejections stem from a perceived obligation to completely identify the structural aspects of a protein marker before it can be used and claimed in an assay.

This misperception arises from a familiarity with assays of an earlier generation that employed single protein markers that were isolated and sequenced and whose biological functions had been elucidated.

Along similar lines, U.S. and European examiners may reject the assay claims for allegedly not being "enabled" (35 U.S.C. §112, ¶1 and Article 83, respectively). As with the above rejections, the examiner might argue that in the absence of complete structural and functional character-

istics, the applicant has not enabled practitioners to make and use a biomarker.

Moreover, under current views of obviousness and inventiveness, U.S. and European examiners likely will argue that the new biomarkers are not patentable over the prior art. The examiners in this regard may argue that it would have been obvious to apply proteomic profiling technologies such as SELDI and pattern recognition algorithms to identify new makers for assessing the status of any disease or therapy because such tools have been used previously to identify biomarkers for various disease states.

The current approach to proteomic inventions employed by patent officials makes securing patent protection for the assays underpinning personalized medicine costly and difficult. To provide the requisite incentive for innovators to invest limited resources into developing the assays needed to make personalized medicine a reality, the offices should consider revamping their examination of proteomic inventions.

Adapting Key Markers

As meaningful change is unlikely in the near term, innovators seeking to patent proteomic-based assays should consider adapting their prosecution strategy to the current administrative landscape. In general, applicants should consider postponing the filing date of an application until more information about the biomarkers has been obtained. For example, rather than trying to secure coverage for every marker discovered, applicants should identify those key markers that generate the most compelling data. Likewise, the best marker combinations should be identified and claimed in a "Markush-type" claim. Finally, applicants should consider isolating and characterizing each biomarker of interest, and attempt to identify the structure and function of each of the claimed biomarkers. •

R. Brian McCaslin, M.S., J.D. is an attorney at Foley & Lardner LLP. He can be reached at mccaslin@foley.com.



Addressing Life Sciences' Constantly Growing Data Challenges

With the continued explosion of raw experimental data, increased use of imaging, growing adoption of new data retention practices, and movement to high throughput computational workflows, the deluge of data can overwhelm life science research organizations. To review the options and opportunities for resolving these crushing data issues, please visit www.bluearc.com/html/infocenter

Up Front Inside the Box

The State of Bioinformatics Kits

BRIAN OSBORNE

A toolkit is what to us, the users? It's that collection of software shortcuts that runs a biological idea through the processor. Familiar bioinformatics toolkits include BioJava, BioPerl, BioPython, and BioRuby. Their corresponding languages—Java, Perl, Python, and Ruby—know no science. But a kit knows chem- and bio- or any informatic that presents its face. These languages are not language-like. They are stiff—they reveal an odd love of a standard and change, if at all, with lurching and heavy labor. The kit is what the language is not—a route to an answer through code that reduces and conceals the language. The kit is clay to the language's concrete.

There is some unknown relating the language and its kit. Emit the name of the language “PHP” close to the screen of one of your more inflammable friends and watch in delight as a derision reaction ignites in fumes and sparks in his brain. But that *thing* creates the MediaWiki kit which begat Wikipedia, the Wiki that begat a thousand Wikis. Our most *correct* language, Java, composes BioJava, but it is barbarous Perl that hacks out BioPerl, the kit that draws the most biologists to our fruitful addiction. Is that because PHP and Perl are, in their hearts, kits?

Claim to Fame

People working in and around open source notice that progress is frequently accomplished through heroism, preternatural efforts by the few on behalf of the many (or sometimes by and for the few). The kit is their claim to fame, their playground and their gallery, their blue sky. They weave with words, pulling the bit-net tighter until it's all under their touch. So for that we thank the near-anonymous in this simple way: Sendu Bala and Chris Fields, of BioPerl, and Yaron Koren, from the Semantic Wiki world.

We are nothing if not disproportionate. We place each

leaf in an ontology with care, but then these trees will dot the datascape, far from the massive data mounds. We decipher and notate the genomes but each in our own way, building detailed memorials to the DNA, with no roads between. Connecting all dots is not what we always do. We have created machines that write billions of nucleotides in a day, but will we leave meaning buried beneath? A million scintillants on their dark grids, times the number of phosphorylated flows. That times the number of precious samples and that times the highway, time. We're now like the astronomers with their fourth dimension, measuring genomes in evolution, meta-genomes over seasons, genomes over lifespans. We've hit supernumerary.

Open Source & Ownership

But when you do dig deep in sequence you may find a treasure. A highlight at the recent Advances in Genome Biology and Technology conference (see p. 8) was the talk by a studious Canadian, Marco Marra. Neuroblastoma is a rare cancer, but we all know it because of its frequency in children, the fatalities, and the severity of treatment. Marra and his group wanted the full mutational details of its transcriptome, from cells highly enriched for the tumor initiators, as close to the cold events as you can get. But the expectation and the result were a profound mismatch. What they read was a list of changes in sequences with gene names that are only linked to B cell development, not neurons. The story of this disease is completely transformed by this surprise, the medicine will refocus on new therapies, and hopes of different outcomes.

Somehow if you can sift your haystack your way, you don't worry about the needles. It is not just craft but the feel of collective ownership that pushes open source, a deep wish to create extraordinary function that *all* will use. One current thought is that we may want to wrap a new present, the “Next Generation” toolkit.

The Bio* toolkits were born when “single” was the norm: gene, interaction, protein, message, CPU. We will *now* create the 21st century kit, which will gyre vast hashes about their axes, all ids, terms, and data, cutting, intersecting and jetting off to methods to precisely annotate and detect. Or will we? This next one looks less like a standalone codebase than a knitting together of R and BioConductor and Perl perhaps, or the BioLib project's libraries, a kit of kits.

Can we count you in? There is an ebb and a flow in open source, one worker rests, then you want another to feel the brilliance. The apps are there, the languages are there, we wait to wrap our next gift.

Brian Osborne is a Principle Investigator at The BioTeam. He can be reached at briano@bioteam.net.

Insights | Outlook

Expanding Molecular Diagnostics

LUCY SANNEs

Molecular diagnostics comprise a large segment of the overall *in vitro* diagnostics industry. Worldwide clinical sales in 2008 for kits and reagents sold by diagnostic companies amounted to approximately \$3 billion or more. Additionally, many molecular diagnostic tests available today are not sold as kits or reagents but are laboratory developed tests offered as a testing service by clinical laboratories and/or by diagnostic companies with their own Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory for performing their proprietary tests.

Insight Pharma Reports sat down with James White, CEO of Osmetech (Pasadena, Calif.) to discuss the newer applications of molecular diagnostics, marketplace trends, and considerations for companies in this space.

Insight Pharma Reports (IPR): What are your thoughts on the role of molecular diagnostics and personalized medicine?

White: When you look at the growth rates for molecular diagnostics, in the 15% to 20% range, pharma companies are getting involved. Every new drug they launch has a diagnostic alongside it that has not been factored into the growth rate. What you are seeing is the trend of the pharma personalized medicine numbers adding to the growth rate. The 15% to 20% range is dealing with the traditional end of the market (infectious disease). I think where you are seeing the really exciting growth is with the personalized medicine opportunities. The personalized medicine market will continue to grow with investments from pharma, as well as tests coming from research and tools players. People are looking for easy-to-use platforms that enable testing to happen in hospitals—not just the larger hospitals, but also the medium- and smaller-sized hospitals. The level of investment from pharma companies, in terms of their earlier-stage drugs, has probably been under-called in terms of the excitement we

Further Reading:

The full interview can be found in *Molecular Diagnostics: A Dynamic and Rapidly Broadening Market*, by Lucy J. Sannes, PhD, MBA. Published by Insight Pharma Reports, January 2009. www.insightpharmareports.com/reports/mdx

will see from personalized medicine.

IPR: What are your thoughts regarding other opportunities in the field of molecular diagnostics?

White: You have three big buckets: genetics (inclusive of pharmacogenetics), cancer, and infectious diseases. For us, as a multiplex detection company, there are opportunities like respiratory panels that are becoming more popular. I think that people see them as an exciting opportunity. In addition, there are other multiplex infectious disease opportunities. When you look at the genetic testing market, there is more opportunity there, and the cystic fibrosis market continues to grow. We touched on personalized medicine; warfarin is the poster child, but there will be many more good examples. In terms of cancer, that opportunity could dwarf all the others. I think that some of the cancer tests will become mainstream. It is somewhat early stage, and more complicated in that area, but Genomic Health has done an outstanding job in terms of driving that market opportunity.

We are focused on validating content for our platform and tests that will be mainstream in the hospital. We are looking at the cancer market with interest. Developments like the DxS K-RAS test are very exciting and getting a lot of interest. We are trying to stand back. Rather than being the “umpteenth” company to develop a breast cancer test, we are trying to determine if, as this content becomes properly validated, there are ways that we can partner or work with people in that space. Thus, the big areas (as we see it) are:

- Multiplex detection around infectious diseases. The one- or two-marker tests are doing well now, but opportunity lies with tests like the ResPlex II from Qiagen.
- We are working in genetics with our cystic fibrosis test.
- Personalized medicine, and potentially the biggest area is cancer. The key is to attract validated content and work with winners in the field, putting their tests on our platform.

IPR: What do you see in the future for molecular diagnostics?

White: The molecular diagnostics market continues to evolve, and it is gaining momentum and support over time. Sometimes we are pleasantly surprised at the support from large organizations that you would not have anticipated. There is a continual stampede, from the tools companies entering the market and from the pharma companies, which are far more engaged than they were 12 to 18 months ago. A lot of very good organizations are entering this space and looking at how they can improve patient care and deliver validated content to the market. I think the signs are strong and the segment as a whole is holding up very well, even in these economic times, because of the benefits of personalized medicine. I believe the key things we have to focus on and deliver in this market are validated content, further development on the research side, and driving this content to the market by researchers and tools companies.

Lucy Sannes can be reached at sannes@att.net.

Computational Biology

Keeping Science Connected the Web 2.0 Way

The rise of scientific social networking is paying dividends for users and entrepreneurs alike.

BY ALLISON PROFFITT

If we could get people to work together, we could really change medicine. We could change discovery, we could speed up the rate of treatments, and new targets coming to the market. And it's a little bit more altruistic, I think, than just normal social networking."

Collexis COO Stephen Leicht and many others may be taking advantage of the social networking trend started by MySpace, Facebook, LinkedIn, and their ilk. But in customizing it specifically for scientists, they envision dramatic benefits, from driving research to searching for career opportunities. There are myriad offerings. Some, like ResearcherID and Elsevier's 2Collab, are focused on identifying potential collaborators. Others—Harvard's Catalyst, Hershey (PA) Research Park's KnowledgeMesh—manage the research network and skill sets within a set boundary. And still others are trying to take the best of the web, and make it work for science.

Leicht's project is BiomedExperts.com, launched in April 2008. The community applies mining technology from the parent company, Collexis, to profile documents and experts within organizations. BiomedExperts boasts some 1.8 million profiles in the system based on mining of PubMed literature, and more than 10,000 registered users, 80% of whom are life sciences researchers. When a researcher logs into the application, "we already have a profile for him," Leicht explains. "And if he wants to know why we think he's an expert in ventricular tachycardia, he can see the specific articles



Collexis' NetworkViewer shows a researcher's contacts (blue) and his extended network (yellow).

[that] we've linked to him."

The pre-populated community brings value to the user, Leicht says. "We take the coauthors of each expert and use them as a proxy or starting point for a social network... Even if [a researcher] never joins, I can see 50 people on here that he's connected to, so I've got a preliminary social network for [him] based on the people that he has done research with over the course of his career... I can connect any Ph.D. student who has written two or three papers within five or six steps to anybody. You can connect them to the head of NIH; you can connect them to the last three people who won Nobel Prizes."

Members can visualize these networks by strength of association (first vs. second level coauthors), chronology, area of interest, and geography. Scientists in the network can message each other, track their connections, and find experts by topic.

BiomedExperts' registered users represent more than 1,800 organizations and institutions from 137 countries, totaling more than 21 million connections made through the site. The site attracts more than 60,000 unique visitors a day. Currently the data for the public network comes from PubMed, but Leicht says that soon the company plans to add grant information and patents to the mined resources.

Invited to the Party

Brian Gilman, founder and CEO of the online network **SciLink**, thinks of his online network—tagline: "Science. Connected."—like this. "Here's the party, we set up the room, we put the drinks out, we set up the stereo and it's up to the scientists to come in."

Similar to BiomedExperts, SciLink also mines the data for publications and connections. When a new user signs up, he or she claims their profile and publications, "and then we show you who you're connected to—auto-magically, if you will," Gilman says.

SciLink members include scientists, clinicians, journalists, venture capitalists, and others interested in the community. "The first community to get in here is biology types," Gilman says, "because I think they're sort of ready. The physics community, we're not offering all of the tools the physics community could use, but they're not as huge as the clinical and bio med community. But there are physicists in here. It's kind of neat."

The site started in August 2007 (see, "SciLink Scours the Web," *Bio•IT World*, Oct 2007). Today, with 10 million profiles in the database mined from literature and 120 million connections, Gilman claims it is the largest scientific social network.

(CONTINUED ON PAGE 18)

Coming Soon!

Biomarker BRIDGE

An interactive collaborative network community for pharmaceutical, clinical and IT professionals who are interested in gaining and sharing knowledge about technologies, research, and regulatory issues around life science.



Join and become a member of the Biomarker Bridge - enrollment opens soon!

First 200 charter members will receive a complimentary subscription to Biomarkers in Medicine

- Create a personal profile
- Browse the membership library
- Complimentary access to CHI Biomarker Events and Webinars*
- Search continually updated Speaker Abstracts White Paper and Poster library
- Network with your peers through Forums, Blogs and Polls
- View continually updated Biomarker News
- Receive on-line access to life sciences publications

* Membership level determines access

Membership Options:

There are several different types of membership access to the Biomarker Bridge. Members to the Biomarker Bridge are granted access to a password-protected web site filled with information, tools and files specifically for Biomarker Bridge members.

Interact with your peers and engage in member-driven research, presentations, webinars and more.

The *BEST* decision you will make in 2009

www.biomarkerbridge.com

If you have questions, comments or suggestions for the Biomarker Bridge, email:
biomarkerbridge@healthtech.com

Biomedical Research Initiative Discussion Groups for Excellence



Cambridge Healthtech Institute

Computational Biology

Science Connection

(CONTINUED FROM PAGE 16)

"We definitely want to foster the kind of interaction [a researcher needs when] thinking about their next great idea. People spend a good deal of time on the site. It's fully customizable to them. We do a lot of really fun Web 2.0-y things... and it's about networking, it's about connecting with people."

Asset Management

Epernicus, another scientific social networking site, aims to connect researchers with "very specific expertise or skills" within their networks, says Vivek Murthy. "[The four founders] were researchers who actually decided to create this platform to solve a problem that we and our colleagues had," says Murthy. "We brought to it a very intimate understanding of science, scientific networks, and how scientists think and what their workflows are like. A lot of that familiarity with the scientists is reflected in the design and the interface itself."

Epernicus uses automated networking rather differently than SciLink and BiomedExperts. Instead of mapping publications and author relationships, Epernicus connects people based on location and advisor relationships. "If I join and I indicate that I'm in the Department of Microbiology at Brigham & Women's Hospital that fits within **Harvard** University, and that my advisor is Jack Smith, my network will automatically pull all of the people that are advised by Jack Smith within the Department of Microbiology, within the Brigham, within Harvard," Murthy explains.

Epernicus' real distinguishing feature is its asset management tools. Assets "are the topics and materials and methods with which a person has expertise," Murthy explains. "They're self-designated items, but they give people an understanding of another scientist's skill set. That, unfortunately, is not information that is easily gleaned if you just look at their publications, or their LinkedIn profile, or their resume even. A lot of it has to

do very specifically with their experience has been, and that's not easily captured." The asset cataloging makes the search feature in Epernicus quite powerful, Murthy says.

Epernicus rolled out their beta solution last summer. Currently the network hosts 10,000 members, and, based on the company's own research, claims to have the largest representation of researchers at top institutions including Harvard, **MIT**, Berkeley, **Stanford**, and **UCSF**.

Selling the Network

All of the networks seem to agree on keeping the networks free for users, but that doesn't help the bottom line. The

An Epernicus researcher profile lets a researcher present his or her resume as well as give details about his or her research.

audience for these types of applications is small, says SciLink's Gilman. Networks are "vying for 2 million people tops," and competing with the non science-specific offerings of Plaxo, LinkedIn, and more.

"There are a large number of smaller players," Epernicus' Murthy says. "A lot of these players have changed their business models and offerings over time." With a web-based product it's easy to adjust.

"We made some assumptions about what they might want, but we're learning from our users every day," Gilman says. "That's why we started this site, because we want this for scientists, by scientists, and so that they can represent themselves using their own language."

There's the altruism Leicht was talking about. But Gilman is acutely aware of the bottom line: "The question is, how are we all going to make money?"

For SciLink, the answer for now is to focus on the job market. "Many people are doing an internal assessment right

now," Gilman says, figuring out what their skill set is, and how and where they are represented on the Web. Gilman is seeing SciLink being increasingly used for recruiting and job hunting. "There's no cost to the user," he says. Recruiters pay to get into the network and hiring managers search the membership. Two of SciLink's full time employees are dedicated to facilitating the job search offerings, and the company did break even in 2008.

An alternative option to making the "altruistic" venture profitable is to sell the solution to companies and other groups who want their own network. Collexis' mining technology, upon which BiomedExperts is based, is already in use. For example, **Johns Hopkins University** uses it to profile 3,000 researchers in its schools of medicine, nursing, and public health. The state of South Carolina uses the technology in its Health Sciences South Carolina initiative to link the half-dozen largest hospitals and universities in the state. Leicht estimates that Collexis does custom expertise profiling for 25 to 30 of the nation's academic medical centers.

SciLink is also selling a private channel package, using the Amazon cloud. Within the custom-created white labeled solutions, SciLink and the institution own the customer jointly.

Epernicus hopes to capitalize on the same need, and is running a pilot program with **Genzyme**. Murthy recognizes that pharma and biotech companies have internal networking needs and would like to do many of the things for their internal scientists that are currently featured in Epernicus, such as "helping them understand each other's skills, helping them communicate through tools like Bench Q, which is a communication tool we have, helping people understand each other's assets and search efficiently."

Epernicus is also "considering partnerships" with companies that want to serve scientists, "lab supply companies, recruiting firms," but the company isn't looking to on-site advertising. "Our goal is to create a really clean experience for people on the site," he says. Altruism again. •

Computational Development

Chasing Cellular Systems Biology

Cellumen uses high content screening to flag tox-causing compounds.

BY JOHN RUSSELL

Founded in 2004 and based in Pittsburgh, Cellumen is one of a small but growing number of biotechs working to use high content screening in cell-based assays to identify promising drug candidates and to red-flag troublesome tox-causing compounds.

It will be interesting to watch the march forward of imaging based methods. No less an authority than Eugene Myers, inventor of BLAST, has said he thinks the amount of imaging-based data will eventually overwhelm even the genomics data flood (See, “[Imaging Informatics](#),” *Bio•IT World*, Jan/Feb 2009).

What sets Cellumen apart, says Don Taylor, director of marketing and son of Cellumen founder Donald Lansing Taylor, is its ‘cellular systems biology’ approach.

The company labels its approach as Cellular Systems Biology. Like several of its counterparts, Cellumen treats cells with compounds, uses a fluorescent reporter system to monitor molecular activities, and develops “fingerprints” which may be associated with disease, toxicology, or health. Two major differentiators, according to Taylor, are its proprietary “gene-activation” systems, which enable effective targeting of entities to be tracked, and Cellumen’s ability to measure and readout many parameters in a single assay.

“The driver [for starting the company] was to be able to address the more advanced cell-based high content screening assays that would be required to meet the new demands in drug discovery and drug development,” he says. Cellumen’s brand of cellular systems biology currently involves using single cell types in multiplexed cell-based assays which can measure six or more parameters per cellular screen, according to Taylor.

The company offers both products and services. Early on, the focus was somewhat more on drug discovery. Today, the emphasis has shifted more to tox

screening. The idea, of course, is to identify problems (or promising candidates) much earlier in the drug discovery process. It has been working primarily with an immortalized HepG2 human cell line and primary rat hepatocytes. Assays are compared with “CellCiphr” profiles, fingerprints Cellumen has already developed by screening a library of compounds, and characterizing them. The revealed biology can be fairly detailed.



“It’s been difficult [to win acceptance] even though there is all this talk about identifying toxic liabilities early on with ‘predictive’ approaches.”

Don Taylor, Cellumen

“In the case of cell models of disease for protein-protein interactions these are actually recombinant proteins that we co-express using adenoviral vector delivery,” explains Taylor. “On each one of those expressed proteins we attach a fluorescent label. One is a GFP (green fluorescent protein) and the other is an RFP (red fluorescent protein). And for

one of the proteins, they will actually attach an anchor to the nucleus. Then the conjugate protein is tagged with a shuttling signal so that basically shuttles back and forth freely between the nucleus and the cytoplasm. When co-expressed, those two proteins are actually connected and they bind together in the nucleus so that you get a signal that is orange.”

“When you start screening for compounds, as soon as there’s a disruption there’s a rapid response so if there’s a disruption in the protein binding the shuttling, the anchor to the nucleus holds but the shuttling component disrupts. It goes into the cytoplasm and then you can very quickly detect the green versus red signal and that’s also a reversible process so that you can wash out the reagent and perform the process over again. So that’s a way to actually, in a high throughput capacity, use high content screening screen for compounds that have this protein disruption capacity.”

Cellumen’s HepG2 panel has ten parameters ranging from oxidative stress to mitochondrial potential, to nuclear morphology and so on. “What’s important is the ability to identify the right biological activity that one is looking to measure and it’s not just a single endpoint,” says Taylor. The entire process is highly automated. Cellumen receives compound from customers in powder or as a solution and produces fingerprints associated with the compounds.

“What I mean by fingerprint is that Cellumen, not only do we measure six or more, 10, even 11 parameters per assay but we also offer a 10-point dose response curve. So for each compound we use 10 concentrations from nanomolar to micromolar. We also measure the compounds across three time points. We have an acute, an early, and a late stage time point that usually is anywhere between one hour, 24 hours and 48 hours or 72 hours, depending on the cell background.” Tay-

(CONTINUED ON PAGE 22)

Computational Development [GUEST COMMENTARY]

Agile IT Strategy in Pharmaceutical R&D

Clearly communicated aims should be the goal.

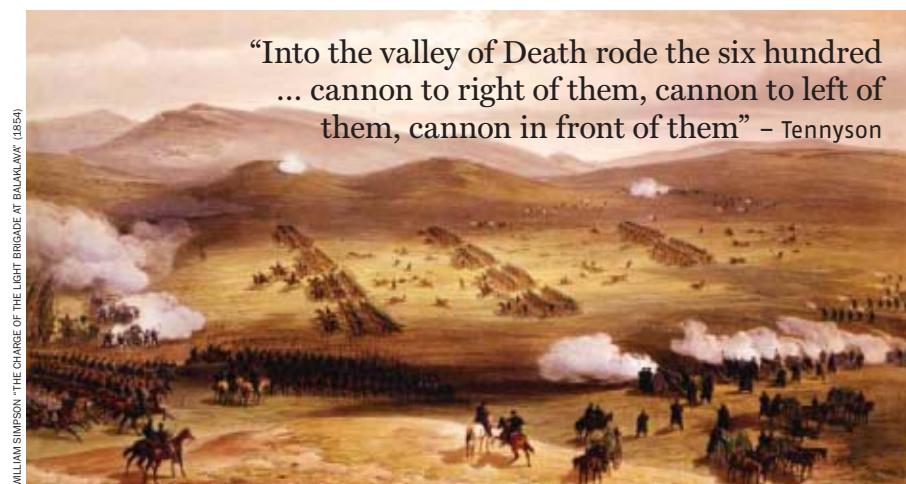
BY ANDREW CHADWICK

As pharma IT groups face tightening budgets and loosening ties between research areas, is an IT strategy spanning multiple research areas still a meaningful aim? To what extent can agile, rapid delivery methods and the increased use of commercial packages ease the burden on internal development, without threatening the coherence and span of the information that scientists need?

Strategy—the “art of war”—is primarily about deploying resources into the right area: fast-moving cavalry to take on local challenges in a flexible and responsive way, while slower-moving forces take strategic objectives over long periods. For IT in drug discovery, a strategic objective is typically to improve creativity and the ability to predict adverse events, which necessitates sharing information between projects and sites. But this will only happen if R&D leaders demand that projects go beyond their own immediate operational needs when recording and organizing their results.

Attempts to define a multi-year, global IT program are typically frustrated by constant change—in research methods, in IT technology standards, and in company structure. Agile development helps to compensate for this. It allows rapid delivery of business benefits by allowing information users and creators to collaborate and find the best means to reach a defined and shared business objective. Rapid prototyping and iterative development methods such as DSDM and Scrum can clarify key requirements quickly, identifying risks in time to overcome them.

However, small, fast-moving teams—the “light horse regiments”—can find it hard to reach, or even identify, objectives that lie beyond the immediate sight of responsible users. Light cavalry alone cannot win a war. The famous mistake during the *Charge of the Light Brigade* was not a failure in strategy but a failure in



The Charge of the Light Brigade (1854) was supposed to attack cannons being moved by the Russians but a mistake resulted in a frontal attack on the main Russian battery at the head of the valley (to left of the picture). Surprisingly, about 500 of the 673 British cavalry survived this charge, helped by a flank attack by French African troops.

communication, the light cavalry charging at the target they could themselves see, instead of the one that their general could see over the hill.

So the longer term aims of a strategy need to be both clear and communicated. What should these aims be for a modern discovery organization?

Capturing Knowledge: Intent and Infrastructure

Sustained improvement in R&D performance requires use of one project’s results, positive or negative, to help improve the conduct and success of future projects. This means pulling information upstream against the flow of the R&D process, which improves the pipeline quality in several ways:

- Clinical chemistry results may carry clues to mechanism or future side-effects, useful to the next discovery project.
- Patterns of activity across multiple targets indicate safety margins.
- Linking targets, pathways, and disease interventions allows repurposing of approved drugs.
- The prevalence of different failure modes should determine what efforts to apply in “de-risking” projects. The most effective plans examine the high-risk

areas first.

- Comparisons of predictions and outcomes calibrate the reliability of assays, *in silico*, and human predictions, which influences their priority and best role in pipeline filtering. This may also help to eliminate tests that add no value, or worse, destroy it.

Such upstream flow requires active intervention by management to set a strategic intent. For example, a knowledge sharing contract will encourage project teams to release and organize information needed by others. Such cross-project knowledge management can meet resistance, but this can be overcome if the mutual benefits are made clear. The key to open information doors between projects is investment in metadata:

- Identify commonalities in pathways and mechanisms to compare projects.
- Pool reliability and calibration information to compare assay effectiveness.
- Agree on data definitions to allow the pooling of results obtained at various places, times, experimental systems.

Complete worldwide standardization of data and IT including use of strict data standards is feasible—indeed essential—in Phase 3 clinical trials, but not realistic in a discovery culture. In our experience,

(CONTINUED ON PAGE 22)

Bio-IT World

OnDemand Webcast

Electronic Submissions – Adobe's Prescription for Insuring Success

In the time-consuming and costly battle to gain market share, faster time to market equals a competitive advantage for life sciences companies. Efficiencies in the product development cycle impact the bottom line. To succeed, companies must streamline the collaborative processes of creating, assembling, and delivering accurate, approvable electronic submissions. Adobe LiveCycle ES (Enterprise Suite) solutions help simplify the submission process, enabling life sciences companies to achieve earlier filing dates, greater accuracy and quality of data, and ultimately faster time to market.

The goal of this webcast is to focus on the creation of submission components, and to demonstrate how Adobe LiveCycle ES software can be applied to realize significant process improvements and thereby gain competitive advantage in the creation and delivery of electronic regulatory submissions.

Automate the creation of regulatory compliant electronic submissions

- Reduce the cost of creating electronic submissions
- Improve document collaboration across your organization
- Satisfy global regulatory requirements
- Improve data capture, information assurance, document output, process management, and content services
- Gain a competitive advantage

Realize immediate Return On Investment

- Reduce/eliminate errors associated with manual processes
- Reduce/eliminate the costs of producing paper
- Insure regulatory compliance with mandates such as E6 and CFR Part 11
- Automate the secure and timely delivery of submission documents to project team members and regulatory agencies
- Impact the bottom line with faster time to market

Speakers:



Kevin Davies, Ph.D.
Editor-in-Chief
Bio-IT World



Maryanne Quinn
Integrated Submission Strategies, LLC
Adobe

Go to "Webcasts" at www.bio-itworld.com

Computational Development

Cellumen

(CONTINUED FROM PAGE 19)

lor adds that turnaround time can be two to four weeks depending on the nature of the screens and the granularity of data and interpretive results requested.

Tox Savings

One practice, which may be unique to Cellumen, is an effort to quantify the savings its tox testing can save. This work is based on a project with consultant [Cambridge Healthtech Associates](#).

Taylor says, "The sensitivity of our test is something that we take very seriously and when we actually quantify this in a financial value proposition. We've been able to use the empirical data that we've received from our Cambridge Healthtech Associates Study, a consortia of over 10 pharmaceutical companies that send Cellumen compounds. We were able to demonstrate using those results that Cellumen projects a cost savings to any one typical big pharma of more than \$91 million per year by applying CellCiphr as a filter at the start of the hit-to-lead phase assuming that one has approximately 400 compounds at the start of the hit-to-lead phase."

Recently the company added a panel to predict cardiac hypertrophy earlier in the drug discovery and development process. "In the past 10 years, nearly 30 percent of all drugs in the United States have been withdrawn due to cardiotoxicity," says Kate Johnson, CSO. "There are many tests available to detect electrophysiological abnormalities of the heart, but Cellumen's Cardiac Hypertrophy Panel is the first systems-based panel to detect other development-limiting toxicities in cardiac cells."

The cardiac tox panel is designed to measure eight distinct toxicity indicators in drug compounds including mitochondrial function, oxidative stress, apoptosis, and cellular hypertrophy and Cellumen says its panel quantifies both the mechanism and time course of toxicity.

Cellumen has worked with several partners—Eli Lilly and Mitsubishi's Tanabe Pharma are two—and last June it announced a research collaboration with the National Center for Toxicology

Research (NCTR). Under the agreement, Cellumen uses its CellCiphr toxicity risk assessment technology to profile blinded samples of known liver toxicity compounds including both failed and marketed drugs for the NCTR. The NCTR will incorporate the knowledge generated by Cellumen to develop a liver toxicity knowledge base. Cellumen will use the profiling data and compound safety data from the collaboration to further develop the diversity in the CellCiphr database and the types of cell panels, as well as to advance the classifier informatics tools.

The company is also collaborating with the Alzheimer's Drug Discovery Foundation. "We're developing biosensors and gene switch cell lines to emulate Alzheimer's disease to allow for the exploration of compounds that may help treat that disease so it incorporates the ability for us to over-express certain proteins coupled with protein-protein interaction biosensors in a complete cellular model."

Despite the progress, Taylor says the competitive landscape is broad: "Probably the most formidable competitor is the status quo. What pharma typically does and what they have been doing forever is a small animal study as primary filter for

any significant toxicity studies. It's been difficult [to win acceptance] even though there is all this talk about identifying toxic liabilities early on with 'predictive' approaches."

"Industry is adopting it slower than they ought to. I don't fault them because there have been many technologies over the past five to ten years that have been touted to be predictive but ultimately did not prove their point of view. So in today's very difficult financial times pharma naturally needs to be very selective in what they choose to license."

But Cellumen is hardly standing still. "We're moving from what we call monolayer cultures, plating HepG2, plating rat primary hepatocytes to human tissue model microarrays, a 3D array if you will. These are going to be created from stem cells, primary cells, and so on. So these will be tissue-specific panels of functional biomarkers that will then go into an augmented CellCiphr classifier that then becomes a predictive tool for human toxicity and so this is to support the FDA's reduce, refine, and replace, the 3Rs Initiative and to help be on the cutting edge of, within ten years, being able to replace animals as the filter for toxicity." •

Agile IT

(CONTINUED FROM PAGE 20)

a federated approach is more practicable. A federated database architecture is a set of heterogeneous databases that, through common metadata, can provide a unified search, allowing greater use of commercial packages to meet local and tactical needs.

There is growing demand from clients to specify and source commercial packages, e.g. instrument data capture, sample management. These days, who builds a chemical search application in house? Use of such packages saves time, reduces risks, and builds the essential user trust in delivery by the IT organization.

However, choice of a commercial package can frustrate efforts toward a single technical architecture or standard (e.g. open source, or service oriented architecture). Unlike manufacturing and financial, there is no dominant 'enterprise research

system'. Overlaps, for example between the scope of electronic lab notebooks and LIMS, can complicate planning and require clarity on the information model and standards for interfacing.

We typically recommend that the available manpower for IT custom build within Discovery groups be focused into two areas: 1) new science that is a source of R&D process innovation and differentiation, such as reproducible processing and objective interpretation of complex biomarker data; and 2) new knowledge that comes from search over multiple results.

Putting value on an IT strategy is easier if there is an end-to-end view. This forges the essential link between science, information, IT, and the business that must fund the IT investment. •

Andrew Chadwick is principal consultant in life sciences at [Tessella](#). He can be reached at: andrew.chadwick@tessella.com.

Handling Next-Generation Sequencing Data

As Bio-IT World has reported with increasing regularity in the past 12 months or so, next-generation sequencing data is producing reams of data—so much that it is producing headaches for platform manufacturers, users, and software suppliers alike. The current iteration of machines from companies such as Illumina, Applied Biosystems (Life Technologies), Helicos Biosciences and Roche/454 can run into gigabases (Gb) of data per day. By the end of this year, some suppliers are predicting sequence throughputs of more than 100 Gb per run, while the cost continues to plummet, possibly to less than \$10,000 per human genome. Extraordinary!

As we have reported (see, "The DNA Data Deluge," and "WikiLIMS—Next-Gen Data Management," *Bio • IT World*, April 2008), the data deluge is affording software and workflow vendors no end of opportunities to engage with core laboratories and other users who, upon plugging in their expensive new sequencer, quickly come to the sobering realization: "What do I do now?"

This supplement features insights from two companies that can help next-gen sequencing users find ways to store, distribute and analyze the torrent of data their new machines have unleashed. Denmark's CLC bio provides next-generation sequencing data analysis software that is proving highly popular in the field. Recent deals with the J. Craig Venter Institute and Illumina are two that have been announced.

San Diego-based Accelrys is not a software supplier traditionally associated with the next-gen sequencing market, but that is about to change. The firm's Pipeline Pilot software (originally developed by SciTegic) is widely used in various workflow applications in academia and big pharma, and is being considered for potential use by sequencing firms including Illumina and Oxford Nanopore Technologies.

The next-gen sequencing market continues to grow beyond even the most optimistic expectations, as more and more applications become amenable to the power of the technology. The firms and products described in this supplement offer worthy solutions to the key challenge underlying this amazing technology—getting a handle on the data.

gcatgcagatcgacgtcgta
tatgacgtactggacttagcacgtatgc
tgtgacgtactggactagctagcacgtatgc
tagcacgtatgcaggatcgacgtcgta
cagtgctatgacgtatgcaggatcgacgtatgc
tcgacgtatgcaggatcgacgtatgc
tgcatgcagatcgactcgtagcatcagtgctatgc
tcagtgtagcagctatgc

ACCELRYS

One thousand complete human genomes is just the beginning

Scientists say that the bottleneck in breakthroughs will shift from sequence production to integrated analysis

Translational Medicine is the practice of applying, or “translating,” otherwise isolated or obscure genomic research into a clinical setting. Often called going from “bench-top to bedside,” life science industry researchers in translational medicine seek to understand how a potential therapy will affect a subset of the population before clinical trials even begin. The goal is to give the medicines only to patients who are likely to benefit from them — and avoid the risk of an “adverse event,” the industry term for a toxic side effect.

Outside of the genomic sequencing of patients, much of the information in pharmaceutical development is already being collected.

Companies already monitor dosing, treatment regimens, and especially responses to therapy, and try to put this information together with genomic analysis to leverage biomarkers. But the collected information resides in “silos,” a common term for pockets of information in an organization that are hidden from or not easily accessed by that organization’s researchers because of IT complexity, access issues, or the need for custom software to do the extraction, analysis, and reporting.

The critical challenge is putting different data together to make better decisions.

As the cost drops for DNA sequencing, scientists are finding new applications for the data, such as completely sequencing the genomes of all patients in a clinical trial. New software approaches are needed, however, to capture value from all the data. “There is already a large and growing amount of sequence data that is resulting from the 1000 Genomes Project, the International Cancer Genome Consortium, and related activities that have outgrown existing analysis tools and approaches,” says Scott Kahn, Chief Information Officer at Illumina, a leading producer of the laboratory instruments accounting for the exponential

growth in sequence data.

Now that so much DNA sequence data are available, the bottleneck is shifting to making sense out of it all.

“For each of the different applications that new sequencing can be used for, there is a very different ‘pipeline’ of bioinformatics analysis required, whether a quantitative tag-counting application, ultra-sensitive mutation detection, or comparative analysis of many whole genomes. Each application type has its own variants too, and different labs implement these applications in a different way. Such ‘polymorphic’ experimentation has to then be reflected in the way the data are managed and analyzed. Thus typically, most of the successful adopters



of these platforms so far have relied upon in-house bioinformatics expertise and locally customizable software,” says Clive Brown, Director of Bioinformatics at Oxford Nanopore Technologies, a developer of new sequencing technology based on single molecule sequencing harnessing modified cellular pores.

The variety of these applications and pipelines has required organizations to invest in significant in-house software development efforts, usually requiring specific expertise in SQL, Perl, Python, R, and SAS. Unfortunately, the brittle nature of these custom solutions means they break when the workflow inevitably evolves. “I am skeptical that the instrument vendors themselves can provide ‘one size fits all’ analysis solutions to cater to the bioinformatics challenges inherent in next-generation sequencing,” continues Clive Brown. “I doubt that software solutions of the required flexibility and

SPECIAL ADVERTISING SECTION

scale can be delivered in a timely enough manner. Typically such software is out of date before it is widely adopted. It has also become apparent that third-party software companies are now applying their considerable muscle to the problems of flexibility, accessibility and scale of bioinformatics analysis."

The need for integration and flexibility is seen as an opportunity for Accelrys, a San Diego-based scientific software company that provides Pipeline Pilot, a leading solution for scientific information management and integration.

"Using Pipeline Pilot, scientists interact directly with all the rich data sources in their organizations without the expense of writing custom software for each workflow," says Jonathan Usuka, Senior Director of Life Sciences at Accelrys.

Comparing the ease of using Pipeline Pilot to playing with LEGO bricks, Dr. Usuka explains that the interface utilizes intuitive drag & drop. "The software works like LEGOs. Each operation you'd like to perform on your data - pulling it out of data silos, performing statistical analysis, writing interactive reports — each one is a single 'block.' These blocks are then assembled into larger structures, with data flowing from one block to another."

"Like creating complicated new things with LEGOs, with Pipeline Pilot, you are only limited by your imagination," adds Dr. Usuka. "Although talking about this powerful



platform like it is a toy might seem trivial, I think it captures the intuitive and creative aspect that got researchers interested in scientific careers in the first place. They are inherently creative people, with very special knowledge. Great things happen when we give them the right tools and allow the flexibility for them to put the tools together in ways not yet imagined."

But behind the intuitive interface is some pretty powerful computer code. Accelrys creates a wrapper around the leading algorithms in life sciences, so all the difficulties

associated with different software, systems, open source and publically available tools, and databases are hidden behind the scene. Scientists are given an integrated environment to develop their workflows. "All the components know how to speak to each other," says Dr. Usuka, "so sequence data can be pulled off instruments, assembled, and aligned with in-house or public sequences. It can also be analyzed with other data sets, such as results from proteomics or gene expression experiments." The company has developed collections of powerful components in imaging and text analytics, so that tissue samples and the latest published research can be integrated in the workflow, resulting in constantly updated reports.

"The benefits of using common analysis components that orchestrate analysis through flexible workflows derives from the successes that this approach has seen in other disciplines," says Illumina's Dr. Kahn.



 **accelrys®**

www.accelrys.com



Next-Generation Sequencing and Diagnostics

Signature Series

CLC BIO

Delivering a flexible and user-friendly enterprise platform with NGS analyses

LC Genomics Server, the enterprise platform, for analyzing and visualizing Next Generation Sequencing data was released by CLC bio in February 2009. It features a modern 3-tier server architecture, designed to provide full flexibility on the client-side and on the database backend, and for integration with other bioinformatics applications.

The user-friendly workflow of CLC bio's enterprise solution allows access by both thin clients, any type of web browsers, as well as thick clients, like any of CLC bio's workbenches. Furthermore, CLC bio's publicly available Software Development Kit allows all customers to interface easily with their own proprietary applications and services.

Unmatched speed

In addition to the various integration and development options, CLC Genomics Server includes a number of pre-installed, high-performance computing enabled Next Generation Sequencing algorithms. These are accelerated by SIMD technology and are also found in the desktop application, CLC Genomics Workbench.

Assistant Professor at Rutgers University, Dr. Todd P Michael, states, "The speed of CLC bio's new algorithm for reference assembly of Next Generation Sequencing data raises the bar to a level currently unmatched by any competitor. When CLC bio continues this impressive rate of development, and eventually also handles SOLiD's Color Space analysis in the same convinc-

ing manner, this could easily become a de facto tool for scientists working with Next Generation Sequencing analysis."

True platform independency

CLC bio develops their Next Generation Sequencing solutions through close collaboration with instrument vendors and genomics centers worldwide, including J. Craig Venter Institute and Beijing Genomics Institute, Shenzhen. CLC bio is the first and only company to offer a comprehensive analysis package which can analyze and visualize data from all major NGS platforms: Illumina's Genome Analyzer, 454 GS flx by Roche, HeliScope by Helicos, and SOLiD by Applied Biosystems—now also including Color Space analysis. Furthermore, all CLC



CLC Genomics Server performs reference assembly, de novo assembly, Digital Gene Expression, and ChIP Sequencing to name but a few features.

bio solutions are cross-platform, running on Mac OS X, Windows, and Linux—including 64bit versions.

"We have worked with the beta-version of CLC Genomics Server for several months now and are impressed with the concepts, flexibility and responsiveness of the CLC customer support and development teams. Given CLC bio's ambitious plans for future development, CLC Genomics Server, combined with CLC Genomics Workbench should provide the enterprise level solution we need," declares Justin Johnson, Bioinformatics Manager at the J. Craig Venter Institute.



Visit www.clcbio.com/genomics to learn more.

Program	Time	Memory	Assembled
Maq	39:37:56 hrs	5 GB	83.14%
Sop	27:52:19 hrs	14 GB	83.05%
CLC NGS Cell	2:14:53 hrs	7 GB	85.03%

The CLC NGS Cell, the high-performance computing solution, performs assembly very fast by using SIMD accelerated algorithms. The speed of assembling 86 million 35 bp reads against the entire human genome (about 1 x coverage) is shown above, comparing CLC NGS Cell with the speed of two of the most popular assembly algorithms for short reads. This benchmark test was performed on a standard desktop computer.



Life at Lilly

real people doing extraordinary things

Lead the next generation of drug discovery informatics

Eli Lilly and Company is a leading, innovation-driven pharmaceutical corporation with a firm commitment to help people live longer, healthier and more active lives by making breakthroughs in medicines and treatments.

To find the next generation of drugs, we know we need to use next generation technologies and employ the best of the biopharmaceutical industry.

The Integrative Computational Sciences (ICS) group at **Lilly Singapore Centre for Drug Discovery (LSCDD)**, provides state-of-the-art computational solutions to enable the global efforts of drug discovery, translational medicine and tailored therapeutics at the post genomic era.

We study the genetics base of complex diseases and develop novel algorithms, data analysis methods, simulation models and tools for drug discovery, biomarkers identification, epigenetics research, population stratification and prediction of dose response.

ICS integrative analyzes solutions, and biological interpretations of complex multi-dimensional data are applied in various therapeutic areas such as Oncology, Diabetes, Neuroscience and Cardiovascular diseases.

As an ICS team member with strong analytical and scientific insight, you will have a global impact on the future of personalized medicine: 'The Right Drug, at The Right Dose for The Right Patient at The Right Time'!

Lilly Singapore is expanding and the ICS group is looking for candidates in the following positions:

- **Bioinformatics Manager**
- **Sr. Statistical Geneticist**
- **Sr. Bioinformatics Scientist**
- **Sr. Cheminformatics Scientist**
- **Principal Statistician**
- **Scientific Liaison**

The successful candidates will work closely with their bioinformatics, statistics and software engineering peers at ICS and will collaborate with biologists, chemists, geneticists and physicians at Eli Lilly. We work closely with Lilly System Biology and Drug Discovery Research teams at Singapore as well as with the Discovery Informatics organization and the Global Discovery Statistics group in Europe and USA.

Minimum requirements:

- Ph.D. in Statistical Genetics/Bioinformatics/Computational Biology/Biostatistics/Biophysics/Cheminformatics or a related discipline
- 3 years' post-graduation experience
- Background in Oncology / Diabetes / Cardiovascular Neuroscience
- Statistics: SAS/R/S-Plus/Partek
- OS: Unix/Linux, Windows
- Databases: Oracle/SQL/mysql
- Scripting: Perl, Shell

Preferred Experience:

- Industry experience in Biotech/Pharmaceutical/Drug Discovery
- Experience in the analysis of large data sets such as microarrays, sequencing, proteomics and imaging data
- Strong publications record
- Demonstrated learning agility
- Excellent communication and multidisciplinary collaborative skills

For more information and online application, please visit www.lscdd.lilly.com.sg/lscdd/careers.

www.lscdd.lilly.com.sg

Lilly

Answers That Matter.

Cover Story

INTEGRATING Clinical and Genomics DATA

With assists from Oracle and InforSense, Dana-Farber's John Quackenbush finds ways to mine vast patient data collections to better understand human cancer.

By Alissa Poh



How does a physicist wind up at the vanguard of translational medicine, bridging genomics, bioinformatics, and IT in an effort to shed light on cancer biology? That's among the chief responsibilities of John Quackenbush, professor of biostatistics and computational biology at the [Dana-Farber Cancer Institute](#) (DFCI) in Boston. An affable personality with slightly greying, shoulder-length hair, Quackenbush is a theoretical physicist by training. He rose to prominence during an eight-year stint at [The Institute for Genomic Research](#) (TIGR), founded by Craig Venter, developing and sharing a range of software tools and databases for microarray analysis (see, "John Q: Life After TIGR").

"People think I love building databases; actually, I hate it," he says from his roomy office. "I'm never going to show up on the front page of the *New York Times* with the headline 'Quackenbush Builds Integrated Database.' If I get there, it'll be because of the discoveries such integrated databases allow me to make. So you could say I want to be able to build the tools that will allow me and everybody else here to show up on the front page of the *Times*."

Whether Quackenbush's efforts become fit to print remains to be seen, but the work he is spearheading at DFCI will likely be felt far and wide in the field of translational medicine.

Bringing Bioinformatics to Cancer

In 2002, Quackenbush was considering leaving TIGR, which had begun focusing on microbial sequencing and annotation, whereas his interest was increasingly turning to the clinical space, beginning with a collaboration with Timothy Yeatman at the [H. Lee Moffitt Cancer Center](#) in Tampa, Florida.

Back then, the trouble facing interdisciplinary scientists like Quackenbush was that no one knew where they fit in the traditional hierarchy of academia.

"We really want you here, we just have to figure out where to put you," Quackenbush would hear. "People were very excited about my work in genomics, but they didn't know quite what to do with the bioinformatics part."

He interviewed at several places, even turning down a position as professor of urology at the Univer-

sity of British Columbia in Vancouver, before accepting an offer from the Department of Biostatistics and Computational Biology at DFCI, moving to Boston in 2005. (He also holds a faculty appointment at Harvard's School of Public Health.)

Quackenbush calls DFCI "one of the most progressive places I've seen in terms of thinking about ways to advance science. And I can honestly say it's the least pathological place I've ever worked." New in Boston and driving to work one morning, he spotted a girl selling lemonade by the roadside. She told him one of her classmates had been treated at "the Farber" and her class was raising money for the Jimmy Fund [DFCI's charity]. The level of community support for and patient involvement in cancer research is incredible, he says.

Before hiring Quackenbush, DFCI had recognized that even with genomics becoming democratized, there were opportunities to do new things that were cross-disciplinary. The institute decided to adopt an entrepreneurial model, with the idea of establishing and awarding five years of start-up financial support to research centers that would work across different departments.

"My message during [interview] presentations was consistently about data integration and its value in propelling science forward, which really resonated here," Quackenbush says. Aside from his own research, he has devoted much of the past three years to building the infrastructure necessary for his other mandate: creating a Center for Cancer Computational Biology at DFCI. That could be viewed as

Cover Story

a service, but “it’s a service to allow me to do the things I want to do,” he says.

Quackenbush reasoned that the success of such a center would require integrating genomic information with clinical data, as one step toward improving cancer diagnosis and tailoring treatment for individual patients. There would, however, be stiff challenges in linking microarray data with not only clinical information but also public archives such as GenBank, OMIM, and HapMap, while ensuring quality control and reliability.

“Web services are all very well, but you’re relying on someone else to maintain the data and not change their protocols,” he explains. “Even for GenBank, where things are supposed to be fairly stable, you frequently see them violating their rules for data entry and standards.” Cloud computing could not possibly work in this space, he adds, given the confidential nature of much of the data. Rather than build a large web services model, Quackenbush elected to integrate all this information in a database unique to DFCI.

Around then, Quackenbush crossed paths with Edie Weller, a senior research scientist in his department, during a faculty meeting. Weller, the lead statistician for multiple myeloma, was trying to merge data from different sources—relational databases, raw text files—a nightmarish and time-consuming process involving many Excel spreadsheets. It was particularly frustrating, when designing whole-genome gene expression studies of chemotherapy response for this disease, that she and her colleagues couldn’t obtain immediate access to data on their own patient samples, even for information as simple as sample storage location.

“I knew there had to be better ways of merging information and allowing investigators direct access to the data,” Weller says. “So although I was initially hesitant to bring it up at the meeting, I finally described how we were linking our data, to John. He looked at me like I was crazy.”

“It was nuts, madness on multiple levels,” Quackenbush recalls. It also clearly illustrated the need for merging different data sources together in cancer research. The multiple myeloma researchers in-

Screenshots from a mock up of the Multiple Myeloma Warehouse ClinicalSense installation feature simulated data sets.

vited him to use their case as a framework for creating a data integration warehouse that could potentially be extended to other types of cancer.

Helping Hands

Oracle, with its expertise in capturing and managing clinical data, came to mind immediately as a potential partner in this data integration venture. “There was no point in reinventing the wheel,” Quackenbush says. After submitting a proposal for one of the enterprise software giant’s commitment grants, he was quickly offered \$1 million spread over two years rather than the three he had requested.

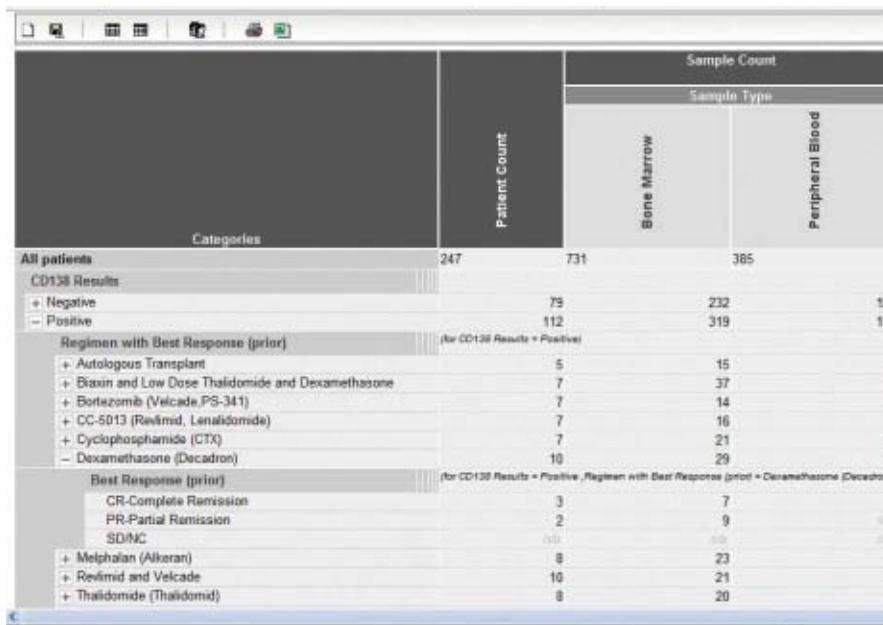
And it was about more than just a financial grant. “We also volunteered technical and subject matter expertise to jump-start Quackenbush’s plans for data integration at DFCI,” says Vijay Pillai, director of strategic planning and business development at Oracle’s health sciences division.

Quackenbush, Weller, and Joseph White, the lead database developer in Quackenbush’s group, attended half-a-dozen workshops over a couple of months, led by Steve Jepsen, Oracle’s senior director for health industries. “We

focused a lot on data security and scalability [during the workshops],” Pillai says. “In moving beyond multiple myeloma, you want to be able to adapt to such growth very dynamically, rather than rebuild your environment. From a security perspective, although clinical studies in different therapeutic areas might reside in one data layer, you want to be sure that the investigators can still only access information they’re authorized to. So we helped Quackenbush’s group think about these design implications.”

The result of these workshops and additional brainstorming was a brand new translational research infrastructure utilizing Oracle’s Healthcare Transaction Base (HTB) and Fusion Middleware components. HTB creates an integrated data repository, whereby researchers access clinical and patient sample information via a single platform and can seamlessly connect this with experimental data. The Fusion Middleware suite, on the other hand, lets them get to their data securely from any location. And a third Oracle component, the BPEL (Business Process Execution Language) Process Manager, allows for safe and, if necessary, multiple transfers of complex clinical data across the entire infrastructure.

At the same time, Oracle hand-picked



The ClinicalSense program gives summary statistics about patient populations in a matrix format that allows clinicians and researchers to drill down into custom queries.

a long-term partner in the intelligent software business—UK-based **InforSense**—to add what Pillai calls their “great visualization capabilities” to the collaboration.

“When you bring different data sources together, you need not just analytics but visualization tools—such as charts and correlation graphs combining thousands of data points—on top of the base integration layer,” he says. “We decided that InforSense’s applications could help keep us completely in sync, in terms of data integration and interpretation.”

Experts at InforSense suggested ClinicalSense, a web-based tool for clinicians and researchers to get summary statistics about patient populations by fashioning row-and-column matrices out of patient sample attributes. “You could construct a query where ‘regimen response’ represents the rows, while ‘sample count’ and ‘sample type’ are chosen for the columns,” explains Mick Correll, InforSense’s director for clinical solutions. “This would result in a cross-tabulation matrix showing the number of available samples broken down by type—tissue or blood, for instance—and grouped according to the patient’s response to a particular regimen.” In other words, users can build more sophisticated queries by defining a hierarchy of attributes, which

then enables them to “drill down” into the results matrix, further stratifying the population.

“It provides, I think, a very rich and interactive web experience that makes the data come alive in the hands of clinicians and researchers,” Correll says.

ClinicalSense leverages InforSense’s next-generation business intelligence platform, besides providing an advanced clinical data model. It’s both easy to use and intuitive, thanks in large part to direct feedback from clinicians throughout the product development process. “The multiple myeloma study at DFCI is precisely the type of problem ClinicalSense was built to solve,” Correll says. “It provided the right balance between out-of-the-box functionality, and flexibility that will enable it to adapt to changing needs in research.”

Institutional Barriers

Like most large-scale collaborations, this one wasn’t without its hitches, particularly with regard to the people responsible for Information Systems (IS) at DFCI. “With these folks, whenever you ask them a question—no matter how benign—their first answer is always ‘No,’ since no access is the most secure access,” Quackenbush

says, only half-jokingly. “The word ‘fragmentation’ has been used to describe this whole problem of having data in different places; I like ‘Balkanization’ instead, because not only are the data being broken apart, there are all these people actively fighting against integrating it. We spent more time and effort negotiating transferring data into this warehouse than we did actually building the warehouse.”

“There was a lot of confusion with IS about the scope of this project; how it fit or conflicted with other IS initiatives,” Weller says. “With how quickly research moves, we felt it was imperative to have individuals who understand the biology, as well as systems aspects, working on the project. Once we discussed this in detail and described our data security model to the IS, things were much easier.”

Nevertheless, Weller adds, such regulatory issues—especially those involving redistribution of data collected from different hospitals—are hardly minor. “I think the time we spent sorting these out will benefit not only our myeloma project, but other initiatives too,” she says.

Upon overcoming these hurdles, the collaborators rolled out their prototype for an integrated data warehouse in May 2008. The warehouse’s full production system has been up and running since November, after two training sessions—one for statisticians, data managers and researchers; the other for clinicians—to teach them the art of accessing and querying the database.

“In both cases, I think the new system was well received, and the feedback we’ve had since has all been very positive,” says Correll, who led the training. “I’m sure modifications will be necessary as it moves forward—this is research, after all—but we’re clearly on the right track.”

Several of Quackenbush’s colleagues, including a few skeptics, were invited to sit in on both sessions. “I remember Beverly Ginsburg-Cooper [senior vice-president for research at DFCI] grabbing me by the sleeve, five minutes into the second session,” he recounts with a smile. “This was after she had watched us go from nothing to a group of patient samples with certain clinical characteristics based on karyotype and trial response, to their gene expression profiles, to a set of genes

Cover Story

John Q: Life After TIGR

John Quackenbush made his foray out of physics in 1992, when he became intrigued by an initiative from the [National Human Genome Research Institute](#) (NHGRI) seeking experts outside biology to work on the Human Genome Project. He spent two years working on the physical map of human chromosome 11 at the Salk Institute, before being hired to set up large-scale sequencing at Stanford's Human Genome Center. When promotion prospects dimmed, Quackenbush headed east to Maryland and Craig Venter's The Institute for Genomic Research (TIGR) in 1997. "The mandate for me at TIGR—going beyond the genome and establishing a microarray laboratory—was really my growing interest," he says.

At TIGR, Quackenbush quickly recognized that there was a woeful lack of tools for collecting, managing, and analyzing the reams of genomic data being amassed. "Our first publication on gene expression in colon cancer included nine arrays, and it was a year's worth of work just to analyze and generate that data," he says. "It really opened my eyes to the challenges and problems with assumptions people have made about biological systems."

Quackenbush recalls early microarray experiments showing that expression of cyclin A1 was a much more appropriate choice of housekeeping gene than the traditional GAPDH, which fluctuated sharply. "What you start to understand," he elaborates, "is that assumptions in biology are often based on little more than gut feelings or historical approaches, and there's nothing better than data to drive a real understanding of what's going on."

Piles of data are essentially worthless without proper management and analysis tools. Given his physics background, however, Quackenbush was comfortable proceeding where most genomicists feared to tread. He continued to write his own data analysis software, creating databases and a variety of open source software tools to help manage the voluminous data being generated at TIGR.

While there, Quackenbush also participated in the scientific workgroup that put forth the MIAME (Minimum Information About a Microarray Experiment) standards. The goal: to allow uniform recording and reporting of microarray data, with the overarching purpose of facilitating the development of databases, public repositories and data analysis tools. It might not be perfect, he says, but these standards have proved handy over the years, especially for finding and correcting errors in published data.

For example, at DFCI he and his colleague Aedin Culhane recently refuted another group's claim to have identified a lung metastasis signature in breast cancer. "Some of the genes they found resonated with us, so we compared their samples with gene signatures in our database and showed that all of the lung metastasis samples fell into the basal-like subtype of breast tumors," he says. "Such tumors are known



to have the highest propensity for metastasizing to the lung. What we recognized, looking at this paper, was that they were really suffering from confounding facts. They weren't predicting lung metastasis; their signature was much more highly predictive of the basal subtype than anything else."

Then and now, Quackenbush's creed is that his software tools must be available in the public domain. "It's my mantra: If you're creating tools, they have to be useful, and they have to be used," he says. "If they're not either useful or used, the overall impact is going to be small; ditto if they're just one but not the other."

Quackenbush was considerably irked, then, when TIGR decided to go with licensing agreements for said tools instead. He remains convinced that attempting to write and market software tools in the genomics space is scarcely a winning proposition. Most of the companies that started out along this path have since gone belly-up. "This was hardly in the spirit of what we were trying to do; we were working to advance the science, rather than create tools," he adds.

This led to his staging—along with two like-minded TIGR colleagues, Steven Salzburg and Owen White—what they humorously called the Open Source Revolution, in 1999. "We decided that if just one of us did it, he'd probably be canned; if all three were involved, [TIGR] couldn't do anything," Quackenbush grins. The trio's efforts to release software to the public domain were mostly welcomed at TIGR, since this eliminated the cost of prying licensing agreements from potential users, and increased the number of successfully funded grants. Nor did it hurt that TIGR was then experiencing a lull from soap-operatic drama, with Venter occupied at [Celera Genomics](#).

But Venter eventually returned, and between the ensuing chaos and TIGR's shifting climate, Quackenbush decided to make his escape. He joined DFCI on March 14, 2005—a date he remembers well for two reasons. His son Adam was born exactly one year later, and March 14, as good geeks know, is also Pi Day.

correlating with response, to PubMed records describing the genes—all as ad hoc queries.” It dawned on the skeptics how this would excite young people doing research at DFCI. “This system presents information in a way they’re comfortable with; they feel invested in the process and better able to participate in data analysis, to see how they can drive things forward,” Quackenbush adds. Ginsburg-Cooper even called it “transformative for research.”

DFCI’s first integrated data warehouse has been constructed architecturally so there’s a path to move forward, Quackenbush says. The institute will pour \$8 million into his cancer computational biology research center over the next five years, which he considers “less than we need, although the center’s built on a model I think we can expand.” He is thus seeking additional funds to accomplish his goal of moving the data integration system beyond multiple myeloma, with breast

cancer as the next likely candidate area.

As well, Quackenbush recently applied for a grant that, if approved, will include funding to create a pilot implementation for data from the Nurses’ Health Study (NHS) at Harvard, the largest and longest-running investigation of factors influencing women’s health. And he has been communicating with multiple myeloma researchers at the University of California, San Francisco, about the possibility of a mirror installation at their end to facilitate data sharing between both groups.

“I don’t know where this will go next, to be honest, but it’s likely to go somewhere,” Quackenbush says. “Our successful collaboration with Oracle and InforSense has put us in a position to think about reaching beyond DFCI and gradually pulling in a lot of Harvard’s multi-institutional spores and their data collections. So there’d be some method to the madness.”

He greatly appreciates how DFCI nurtures the importance of continued research. “You always hear about Harvard eating its young,” he muses. “I came here a little worried that I was going to face all these prima donnas. It’s not to say that there aren’t those with pretty big egos, but there’s a really high level of collaboration here, which is both astonishing and impressive.”

Quackenbush embraced interdisciplinary research early on, and now observes that many scientists in his area of interest are spanning their boundaries, which he definitely endorses. “A computational model is just that; a model plus validation is a discovery,” he remarks. “People are really trying to drive the latter, rather than being held captive to someone else’s experiments. It isn’t true of everyone in the field, but I think it’s an emerging trend; a systems biology approach that is evolving naturally out of genomics and bioinformatics.” •

Today's Industry Challenges Are Tomorrow's Business Opportunities.



Heal, Fuel, Feed The World.

From developing green agricultural production methods to increasing crop yields for a growing population, biotech innovations play a major role in solving the world’s biggest agricultural challenges. Yet current economic realities demand new investment models. Never before has effective partnering and licensing been so critical. Join us in Atlanta for the **2009 BIO International Convention** and access the expert knowledge and resources you need to help deliver on biotech’s promise to feed the world.

Register and reserve housing online today.


Bio
International
Convention
*The Global Event
for Biotechnology*

May 18–21, 2009, Monday–Thursday

Georgia World Congress Center

Atlanta, GA USA

convention.bio.org

A SERVICE OF:

BIO
BIOTECHNOLOGY
INDUSTRY ORGANIZATION

IT/Workflow

Ted Slater's Semantic Technologies

The semantic web doesn't exist, but Pfizer's Slater believes semantic technologies are paving the way.

BY KEVIN DAVIES AND PHILLIPS KUHL

The theoretical benefits of the semantic web for life sciences have been debated for a few years now (see, "Masters of the Semantic Web, *Bio•IT World*, October 2005), but practical examples within pharma remain scarce. Pfizer's Ted Slater is an interesting exception. Slater heads a small group of four informatics scientists in St. Louis called the Indications and Pathways Center of Emphasis (IPCoE), which supports Pfizer research efforts in identifying and validating inflammation targets.

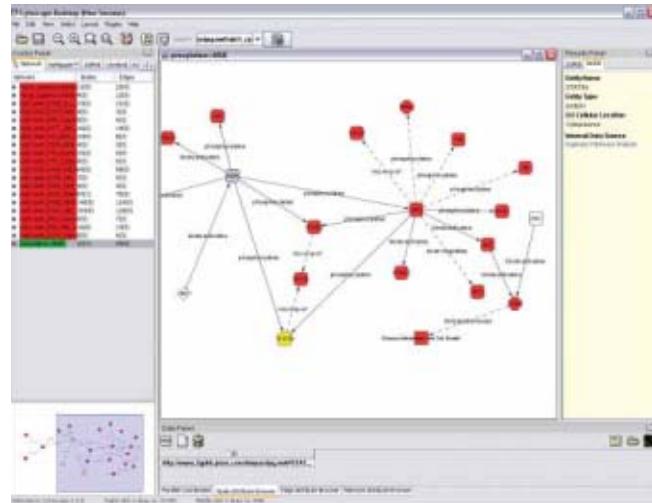
Slater trained as a molecular biologist, but made the mistake of buying his first computer shortly after starting his Ph.D. at UC Riverside. "72 straight hours later, I realized that I'd made a vocational error and maybe I wanted to learn more about computer science instead." He took a Master's in computer science, and went on to work for a string of genomics companies in the '90s, including Sequana Therapeutics, GCG, and Paradigm Genetics. He was the founding vice president for knowledge engineering at Genstruct, joining Pfizer in 2004.

Slater tends not to use the term "semantic web". "There just isn't a semantic web," he says candidly. "It doesn't mean there won't be one in the future, but there isn't one now." He prefers "semantic technologies," so as not to provoke critics who would otherwise argue that he should move on to something else.

At CHI's Bridging Pharma & IT conference last October*, Slater outlined a project focused on pathway data analysis that showed how structuring data in a semantic network could provide substan-

tial benefits over traditional pathways solutions, such as automated hypothesis generation. The effort has gone from idea to reality over the past 6 months.

Using semantic technologies in this way helps to eliminate long-standing problems in informatics like data silos, where information is not interoperable with other necessary information, and data tombs, which simply make informa-



The Pfizer Environment for Knowledge Engineering (PEKE). Knowledge representation refers to triples in RDF and ontologies in OWL (such as BioPAX). The triple store is Oracle 11g's built-in RDF model. Visualization is done with Cytoscape. Content is extracted from the published literature.

tion very difficult to retrieve. Semantic technologies shift the focus from collecting information and making it safe to actually using the information in its proper context to solve research problems. "The computer should be a way of enhancing our own natural ability to reason, in the same way a bicycle enhances our ability to move ourselves around," says Slater.

Be Reasonable

Ideally, Slater says you would like users to be able to reason with the data *in silico*: using "if-then rules," let the computer

generate a hypothesis, then let the scientist decide the potential implications and leverage knowledge to test whether they are supported with data.

"We constantly hear that the Holy Grail is complete data integration," says Slater. "I have bad news—it will never happen! Users are able to set up and start building new, independent repositories of data faster than we can integrate existing data. You will never be able to get it all in one place where it is integrated and useable. The goal instead should be data that are interoperable, even if they are not integrated."

Slater's group helps scientists to study gene expression and signaling pathways in order to identify alternative indications for drugs in development. "There is no easy way to understand what is going on if you look at a list of 1,000 genes that are significantly up- or down-regulated," he says. Even if those genes are mapped onto pathways using a commercial pathways tool, one is forced to work with what amounts to a reference tool. Much information is available on individual genes, but you have to try to tell a story about physiology by painstakingly going through the annotation for each gene one at a time.

An alternative approach lets the computer generate hypotheses based on available data, which would distill the range of possibilities to a few key relationships. If the data are represented correctly, you can use data from disparate databases, such that users can create 'boutique' knowledge bases for their own needs and easily link them together.

Adapting the familiar "triple" semantic RDF format—representing information as a subject, predicate, and object—Slater represents the data as a mathematical

*CHI's Bridging Pharma & IT; Providence, R.I., October 28–30, 2008

graph, with subject and object as nodes and the predicate (the relationship between them) as an edge. One triple's subject can be another triple's object, and so on, until a very large graph of everything known in some domain is created. In this format, the information can be handled with software to build inferences and test hypotheses. His group uses open source ontology development tools to build OWL ontologies and another open-source tool, Cytoscape, to view the data in graph format. For persistent storage, knowledge graphs can be managed in Oracle's built-in RDF data model.

PEKE Performance

One of the goals in data analysis, says Slater, is to use heuristics over the knowledge bases to tell a story. Semantic representations of knowledge allow you to apply expert reasoning to experimental data, which may help explain a particular outcome and in turn suggest a testable

hypothesis. "You don't get inferences in a traditional structured database," says Slater. "We have our share of traditional databases, and we are getting better at data warehousing. For many scientific applications, representing the data as an RDF graph and building for interoperability make the information much more usable. If the description of your problem solution ends with, 'and then the user can query it,' then you haven't thought it through enough." How you structure the information can either lock up the information in a data tomb or set it free.

The experimental system that Slater and his group have developed is called the "Pfizer Environment for Knowledge Engineering", or PEKE. "PEKE handles all of the usual storage and querying capacities of traditional databases, but because of its architecture it has some surprising emergent properties," says Slater. Among these are the ability to create, with just a couple of mouse clicks, new knowledge bases

that essentially automatically interoperate with other PEKE knowledge bases.

Another capability of PEKE is that, because the semantics of each knowledge base are explicit in its OWL ontology rather than implicit in a relational database schema, PEKE supports knowledge bases containing any kind of knowledge with no changes to the architecture. While most PEKE knowledge bases are currently about molecular pathways, the ontology Slater uses to demonstrate how easy it is to create PEKE knowledge bases is the OWL pizza ontology from Stanford's Protégé Team.

Slater says, "PEKE is world-class stuff. We think we can now build knowledge bases faster and cheaper than anyone else in the industry, and do much more with them once they're built." We may still be waiting on the semantic web, but semantic technologies are already paving the way for the next wave of informatics innovation. •



Visit www.eCliniqua.com today!
Subscribe to the eCliniqua eNewsletter today!



CAMBRIDGE HEALTHTECH MEDIA GROUP

Your Connection to Clinical Research

Introducing *eCliniqua* – a new web site focusing on innovative management in clinical trials covering specific topic areas including:

- Trial design
- Data capture
- Regulatory & safety
- Market insights
- Technology solutions
- Other clinical-related critical topics

The web site is the *expanded face* of the *eCliniqua* eNewsletter, one of the most respected sources for news covering clinical research and trials. Both *eCliniqua* mediums provide life science professionals with critical and dynamic coverage on the clinical research enterprise.

Stay informed. Remain competitive.

Bio-ITWorld.com • CHIMediaGroup.com

eCliniqua—Innovative management in clinical trials

Platforms for Personalized Medicine

Integrating clinical information and R&D data will ensure success.

BY NEIL DE CRESCENZO

Personalized medicine promises to improve understanding of the mechanisms of disease and permit more effective patient care. It also stands to transform the focus of the pharmaceutical industry as the cost of drug discovery exceeds \$1 billion, and established blockbuster drugs lose patent protection.

Personalized treatments based on an individual's genetic profile would usually be targeted toward sub-segments of the population, creating new opportunities for drugs that might have failed traditional broad-scale clinical testing to potentially treat target populations. Although it holds great promise, realizing the benefits of personalized medicine presents significant challenges for life sciences companies, particularly surrounding the integration of research and clinical data—both within and between life sciences companies and health care organizations.

The breadth and depth of data available electronically from health care providers, pathology labs, genetic diagnostic labs, and other research institutions is exploding. The challenge is to organize and integrate this clinical information with R&D data to guide research—as valuable insight can emerge when phenotypic data are combined with pathology and genetic data in the context of a specific disease. To ensure an economical and scalable personalized medicine model, pharma R&D will have to link more closely to clinical care delivery. Informatics—from the perspective of the health care organization and pharma—is at the forefront of surmounting this challenge.

Clinical R&D data are fragmented across many silos. As they engage with academic medical centers (AMCs) to develop new personalized treatments, life sciences companies are looking toward standards-based platforms that facilitate integration of internal and external data

sources. Such a platform can provide “vertical” informatics capabilities around specific drugs (from discovery, through clinical trials, to post-market surveillance) as well as “horizontal” informatics capabilities required for portfolio management across the pipeline. Much of the data required to support informatics around costs and outcomes lie within provider organizations. Hence, the architecture for data integration for life sciences com-

The drive to implement informatics capabilities will change the way companies use R&D information.

panies must support internal R&D data models, as well as standards-based data models within provider organizations.

The most ambitious frontier in R&D informatics focuses on the capture and use of scientific knowledge within an enterprise. While companies routinely consolidate drug pipelines through mergers and acquisitions, the knowledge and insight within the merging companies have been difficult to consolidate and leverage. Web 2.0 tools promise to simplify semantic information sharing in an enterprise. Ontology-based search engines combined with natural language processing capabilities can help researchers find and correlate relevant scientific insights buried across multiple data sources, taking content management and data mining to a whole new level.

Health Care Informatics

As diagnostics and therapeutics become more closely tied in the personalized medicine world, the resulting informat-

ics needs are becoming more pervasive within health care organizations. There are three levels of informatics capabilities that health care organizations require to support the transition to personalized medicine:

- **In-silo Informatics**, which focuses on clinical specialties, such as oncology. These capabilities integrate specialized clinical data with relevant content to improve the evidence base for the specialty and guide specific care protocols.
- **Cross-silo Informatics**, which supports disease management and enables health care organizations to track outcomes. These capabilities are most valuable for primary care, emergency care, and chronic disease management, as well as infection control and public health.
- **Performance and Cost Informatics**, which enable cost-benefit analysis of personalized treatments. With targeted treatments serving smaller patient populations, current pay-for-service reimbursement models may prove to be ineffective.

Integrating the applications and data required to support the informatics required for personalized medicine is a challenge for many life sciences companies. Mergers and acquisitions further complicate the IT landscape by introducing more heterogeneity. Over the next decade, the drive to implement informatics capabilities will surely lead to a major overhaul in the way life sciences companies manage and utilize R&D information. Organizations leading the charge must be careful to ensure that incremental investments in applications and infrastructure also provide direct business value along the way, while adhering to a long-term road map. Those that succeed stand to transform the treatment of disease while gaining new levels of productivity within their own R&D programs. •

Neil de Crescenzo is senior VP and General Manager, Oracle Health Sciences. He can be reached at neil.de.crescenzo@oracle.com.



The Path to Drug Development, Regulations, and Industry Trends

PAREXEL's Bio/Pharmaceutical R&D Statistical Sourcebook 2008/2009

The leading and most widely read compendium of pharma/biotech R&D trends, data, market intelligence, articles, and graphs providing fresh insights into the developments reshaping pharma R&D and the drug development industry. Also available in electronic format.



The U.S. Drug Approval Trends and Yearbook 2008/2009

A one-of-a-kind compendium filled with hundreds of key trends and metrics used by professionals to gain industry insights, new benchmarks, and analysis to help plan their own R&D projects, improve company performance, and assess various drug approval options and strategies.

New Drug Development: A Regulatory Overview 2008

Addresses the most cutting-edge developments redefining how new drugs are developed and regulated today. The book is considered the "go-to" resource for regulatory, clinical, project management, training, and other drug development disciplines navigating the FDA's drug development approval processes.

Expediting Drug and Biologics Development: A Strategic Approach

Using a unique "reverse-engineering" approach, dozens of leading experts with extensive experience in all disciplines of drug and biologic development show how careful planning and a sharp focus on the end-goals can be used to expedite the most complex product development program.

Biologics Development: A Regulatory Overview

The only text book on biologics development and regulation in the post-CBER/CDER Consolidation Era! Offering an expansive examination of the FDA's regulation of biologic products, from preclinical testing to post-marketing regulatory requirements, and from user fees to electronic submissions.

Regulatory Reporter Newsletter

A monthly newsletter providing regulatory information for the pharmaceutical, medical, and allied industries. Annual subscriptions available.

- Emerging pharma/biotech R&D trends
- R&D performance and productivity
- Phases of clinical development
- Informed consent & IRB review

Purchase your copy(ies) today for your colleagues, research staff members, and professionals involved in the management and conduct of clinical research.

To order, call 1-800-856-2556, visit the "Publications" section on www.barnettinternational.com, or complete the Publications Order Form on page 88.

- Integrated approach to data management
- Nonclinical drug testing
- The Biological License Application (BLA)
- Priority drug review destinations
- CDER and the NDA review process
- Pediatric studies initiative
- Post-approval changes to marketed drugs
- Phase 4 commitments

Barnett International: A division of Cambridge Healthtech Institute, www.barnettinternational.com
250 First Avenue • Suite 300 • Needham, MA 02494 USA • Phone: (800) 856-2556



www.barnettinternational.com



Cambridge Healthtech Institute's Seventh Annual

Bio-IT World

CONFERENCE & EXPO '09

April 27 – 29, 2009 • World Trade Center • Boston, MA



Enabling Technology. Leveraging Data. Transforming Medicine.

KEYNOTE PANEL:

The Future of Personal Genomics:

A special plenary panel discussion featuring...

Robert C. Green, M.D., M.P.H., Professor of Neurology, Genetics and Epidemiology, Boston University School of Medicine and Public Health

John Halamka, M.D., M.S., CIO, Harvard Medical School

Clifford Reid, Ph.D., Chief Executive Officer, Complete Genomics, Inc.

Dietrich Stephan, Ph.D., Co-founder and Chief Science Officer, Navigenics, Inc.

KEYNOTE PRESENTATIONS BY:



Chris Dagdigian, Founding Partner and Director of Technology, BioTeam, Inc.



Clifford Reid, Ph.D., Chief Executive Officer, Complete Genomics, Inc.



Eric E. Schadt, Ph.D., Executive Scientific Director, Genetics, Rosetta Inpharmatics/Merck Research Labs

Event Features:

- Access All Six Tracks for One Price
- Network with 1,500+ Attendees
- Hear 100+ Technology and Scientific Presentations
- Choose from Multiple Pre-conference Workshops
- Attend Bio-IT World's Best Practices Awards
- Connect with Attendees Using CHI's Intro-Net
- Participate in the Poster Competition
- See the Winners of the following 2009 Awards:

Benjamin Franklin
Best of Show
Best Practices

- View Novel Technologies and Solutions in the Expansive Exhibit Hall
- And Much More!

CONCURRENT TRACKS:

Track 1: IT Infrastructure & Operations

Track 2: Informatics for Genomic Medicine

Track 3: Predictive and *in silico* Science

Track 4: Computational Chemistry

Track 5: Clinical and Medical Informatics

Track 6: eHealth Solutions **NEW**

Please mention keycode
U02 when registering!

Organized & Managed by:

Cambridge Healthtech Institute
250 First Avenue, Suite 300, Needham, MA 02494
Phone: 781-972-5400 • Fax: 781-972-5425
Toll-free in the U.S. 888-999-6288

www.Bio-ITWorldExpo.com

Featured Presentations by:



Platinum Sponsors:



Gold Sponsors:



C4 Systems



Bronze Sponsors:



Official Publication:



Track 6:

eHealth Solutions

Integrating life sciences, IT and general healthcare to support the care delivery process, clinical decision support, patient-delivery, and innovative R&D of next generation health IT and personalized medicine solutions.

NEW



Featured Speakers

Austin Tanney, Ph.D., Scientific Liaison Manager, Almac Diagnostics
 Belinda Chen, Ph.D., Deputy Director, IDEAS, Institute for Information Industry
 Dietrich Stephan, Ph.D., Co-founder and Chief Science Officer of Navigenics
 Jon Schoonmaker, Chief, Technical Programs and Operations, SAFE-BioPharma Association
 Ken Buetow, Ph.D. Associate Director for Bioinformatics and Information Technology, National Cancer Institute
 Mark Boguski, M.D., Ph.D., Founder, Resounding Health Incorporated
 Paul Mattes, M.B.A., Director, Life Sciences, Health Solutions Group, Microsoft Corporation
 Robert Lundsten, Director of Biorepository, Feinstein Institute for Medical Research
 Samuel (Sandy) Aronson, Executive Director of IT, Center for Personalized Genetic Medicine, Partners HealthCare
 Subha Madhavan, Ph.D., Director, Clinical Research Informatics, Lombardi Cancer Center, Georgetown University
 Sujith Eramangalath, M.Sc, Senior Analyst & Team Lead, Healthcare, Frost & Sullivan
 Tibor van Rooij, Bioinformatics Director, Génome Québec & Montreal Heart Institute Pharmacogenomics Centre
 Werner Ceusters, M.D., Director, Ontology Research Group, NYS Center of Excellence in Bioinformatics & Life Sciences

Featured Presentations

- Global Trends in Convergence of Electronic Health Records and Clinical Trials
- Scalable Architecture for Genetics Enabling Electronic Health Records
- Online Health Information Retrieval by Consumers and the Challenge of Personal Genomics
- The Logistics of Adaptive Trial Design Sponsored by ALMAC
for Biomarker Discovery and Validation
- Demonstrating the Ability for Health Information Exchanges to Effectively Share and Protect Public Health Information with Necessary Relying Parties
- caBIG®: An Interoperable IT Framework for Personalized Medicine Featured by caBIG® cancer Biomedical Informatics Grid®
- Do Healthcare IT Standards Hamper the Advance of Science?
- Optimizing Patient Care by Incorporating PGx into the Clinic: Testing and Guidance
- Elevating the Research Enterprise - Overcoming the Challenges to Discoverability in Healthcare and Life Science Research Sponsored by Microsoft
- Integrating Personal Genomics into Medical Practice
- G-DOC: A Framework for Personalized Medicine
- Establishing the Collaboration Center of Health Information Application (CCHIA) to Enable Public Health Policy Making
- Service Oriented Architecture – A New Data Sharing Architecture

Pre-Conference Workshops

Monday, April 27

Half Day Workshops

Workshop 1: Application of Wikis by Lifescience Organizations (12:30-4:00pm)
 Workshop 2: Advances in Drug Safety Informatics (12:30-4:00pm)
 Workshop 3: Visualization (8:00-11:30am)
 Workshop 4: Recent Advances in Molecular Dynamics: Target Elucidation and Ligand Docking (12:30-4:00pm)

Workshop 6: Designing Storage Architectures Sponsored by that Address the Explosion of Data in Life Sciences Research (8:00-11:30am)

Full Day Workshop

Workshop 5: High Performance Computing & Storage: Trends and Applications (8:00am-4:00pm)
 John Halamka, M.D., M.S., CIO, Harvard Medical School
 Jacob Farmer, Chief Technology Officer, Cambridge Computer
 John Dey, UNIX Operations Manager, Rosetta Inpharmatics, LLC
 Reece Hart, Scientific Manager, Research Computing, Genentech
 Rick Franckowiak, Information Technology Director, Johnson & Johnson
 Andrew Kaczorek, Eli Lilly and Company
 Giles Day, Senior Director BBC Informatics, Pfizer Biotherapeutics & Bioinnovation Center
 Stephan Bour, Ph.D., Chief Information Officer, Director, Center for Biomedical Informatics National Heart, Lung, and Blood Institute (NHLBI)



www.Bio-ITWorldExpo.com

SAVE THE DATES!

EUROPE

CHI Cambridge Healthtech Institute's Inaugural

Bio-IT World

CONFERENCE & EXPO '09

6-8 October 2009 • Hannover Fairgrounds • Hannover, Germany



Co-Located with BIOTECHNICA 2009,
 Europe's No.1 Event in
 Biotechnology and Life Sciences

www.Bio-ITWorldExpo.com/Europe

Clinical Research [IN CONVERSATION]

A Web MD: Paul Bleicher Reflects on Phase Forward

Bleicher talks about Phase Forward's acquisitions, milestones, and what's next.

*Last November, Paul Bleicher stepped down as chairman of **Phase Forward**, a leading e-clinical software company that he founded more than a decade ago. Bleicher remains a director of Phase Forward, but these days can be found occupying a small office in a new health care informatics start up, where he serves as chief medical officer. (Details of that venture should emerge in due course.)*

Kevin Davies caught up with Bleicher, who reflected on a decade at the helm of a leading player in clinical technology. (This is a slightly edited version of Bleicher's remarks: see the full video interview at: www.bio-itworld.com/lsw/paul_bleicher)

Bio-IT World: You trained as an oncologist and immunologist.

What did you see about the role of the Internet in clinical trials back in the mid 90s?

Bleicher: I actually took a long journey out of Harvard Medical School... and found my way into the pharmaceutical / biotechnology industry, because I thought that was the place where progress in basic science got translated into progress for people and for disease mitigation. At that time, as I began to get into it, both as an investigator at Mass General and then at a CRO and at a biotechnology organization, I learned that there was a lot of inefficiency. A lot of things had been tried involving laptop computers, even large computers, but with the emergence of the Internet and the ability to be able to view this through a web browser, I saw a new opportunity and was at the right place at the right time.

What was the original mission of Phase Forward when you launched in 1997?

The mission was to solve two of the problems I thought had plagued me in the various clinical trials I'd run before. First, we needed to get real-time clinical trial data ... and make that actionable, both from a safety perspective and from all perspectives, making sure that data was



in house and high quality as soon as possible, which was consistent with the best practices of science.

The second was all the operational information—about who did what, when, where—was so important to someone who was running a clinical trial, to make sure these resources were appropriately spent and managed and that the trial could be done on time. It was those two things that I saw were the opportunity and became our mission at Phase Forward—to create a high quality, regulatory compliant, complete application for the management of clinical data and for the management of the operational aspects....

Has the industry complied or are there unresolved issues?

A little of both; obviously, that's not unexpected... Certainly one always dreams as an entrepreneur to be highly successful, and it did to my wildest expectations. It started a trend that changed the industry. It's something that's been adopted by a large percentage of pharmaceutical companies, and is in use in at least half of the clinical trials by best estimates... On the other hand, a lot of the efficiency that can be obtained... both from operational aspects and from a variety of process changes, etc. haven't been perfected at this point. They haven't gotten to the point where we as an industry have made best use of them.

As you say, 50% of trials are still not performed electronically.

Where is the problem?

I actually don't think it's a problem at all. It's following a standard technology adoption curve that's been known since the 1950s and was popularized in Crossing the Chasm... but the pace at which different industries do technology adoption is different. The pharmaceutical industry clearly, a number of years ago, crossed the chasm... and adoption at the 50% level is at the peak of the technology adoption curve... Compare that to the industry I'm working in now—electronic health records. That started at a very, very low percentage, very similar to where clinical trials were in 1997. By best estimate is in the 12-14% range, as compared to the 50% range [for clinical trials], so I'd say we're quite successful in that regard. Some companies are still waiting for more feedback, more evidence of successes from

(CONTINUED ON PAGE 44)



We've Made Getting it Done Right a Science

At ERT, we're passionate about enhancing your clinical trial and drug approval efforts. Everything we do has one focus. Delivering reliable results to empower YOUR expert decisions and quality outcomes – each and every time you work with us.

- Hands-on Experts
- Unparalleled Service
- Proven Process
- Leading Technology

When you need to get it done right, put our experience to work for you.

+1 866 538 2808 | www.ert.com



Cardiac Safety
Solutions



EDC Solutions



ePRO Solutions



Clinical Research
Consulting

Clinical Research

MacGarvey Establishes Quantitate's U.S. Operation

The Cambridge office expands reach of clinical trial services and products in North America.

BY KEVIN DAVIES

Andrew MacGarvey, the newly installed president of Quantitate's American office in Cambridge, Mass., has a mission: to expand clinical trial service capabilities in North America, while continuing to manage global sales. Quantitate was formed 18 months ago from the merger of two 15-year-old European biometric CROs—Oxford Pharmaceutical Sciences and Statwood. "The heritage is really statistics and programming," says MacGarvey. The firm is a biometric CRO, with more than 200 personnel. "We are a niche specialist provider, but we have very large critical mass," he says. "The move to establishing a U.S. headquarters is in response to the terrific reception of our products by companies here in the States."

The firm has two major groups of customers. One group seeks Quantitate's expertise with statistics, study design, programming expertise, and data management, which is married with a clinical partner. Quantitate is also picking up FSP (functional service provision) work for customers, providing up to 40 heads on a specific project, whether it be programming, statistics, or medical writing.

The company has offices in the U.K., Poland, and South Africa. The U.S. office began as a sales office, but MacGarvey's charge is to build up U.S. operations, leading a small team that currently has three business development staff and a handful of programming talent. "I've come over to expand the operational capability," he says. "We've already got U.S. customers working with us, but some of our U.S. customers were saying you need to have more operational people over her."

Quantitate's best known clients include top-five pharma companies, but MacGarvey has his eye on mid-size companies and biotechs. "We want to keep a

variety of work for our workforce, keep the retention rates high," he says candidly. "If you restrict yourself to larger customers, the work can be a little repetitive. Also, from a strategic point of view, it helps to have these emerging companies and building relationships with them earlier. We hope to grow with them."

Quantitate provides services in clinical data management, programming, statistics, and medical writing. It currently partners with clinical companies chosen based on the therapeutic area or regional



The move to establishing a U.S. headquarters is in response to the terrific reception of our products by companies here in the States."

Andrew MacGarvey, Quantitate

considerations. But he wants to bring clinical services on board and become full service. Among the companies that MacGarvey comes up against are i3, ICON, and Kendall.

Virtual Reality

A lot of emerging prospective clients are almost virtual, MacGarvey observes. "They're getting funding to get these studies kicked off, but they tend to have very few operational people." Quantitate takes a very consultative approach to business development, he says. "We don't go in with a slide deck necessarily. We'll go in and have a conversation, talk to the person. They want to know right down to the study design what our advice would be, and which vendors to work with to help

put the study together."

MacGarvey's been advising "a local company" about their data strategy, trial management systems, and EDC, to help them formulate a data strategy. "The larger companies tend to have more bodies on the ground... but they'll be interested in talking to us about European [trials]... If they're going to submit across the Atlantic, they'll want to talk about expediting European submissions."

Quantitate works with Phase I onwards. Typically, the client will have preclinical data. "Sometimes they'll have Phase I studies run already, and we'll come in to look at the Phase II trials. Because we are biometrics... we're being used by lots of different groups."

Easy EDC

Quantitate offers its own SAS-based EDC system called ClinNav, but MacGarvey says recommendations are determined by the needs of the customer. "The Clin-

Nav system is very simple to use, without lots of bells and whistles... You capture the data, you can clean the data." When one client wanted to launch a quick trial at various sites around the world, MacGarvey's advice was: "Look, you're going to want something that's been used in a lot of the sites already, because you're trying to go so quickly. So you're probably better off using something like the Phase Forward product." As simple as ClinNav might be, it still requires site training, which takes time. The company works with Oracle as well.

MacGarvey said his team is seeing "a great deal of take-up" in adaptive trials. "It's becoming a very hot topic back in Europe, and I guess it's the same here."

(CONTINUED ON PAGE 44)

LIFE SCIENCE REPORTS YOU CAN RELY ON

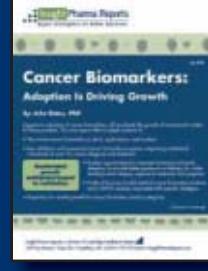
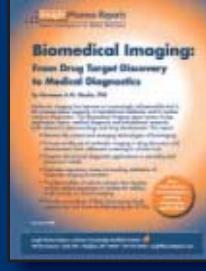
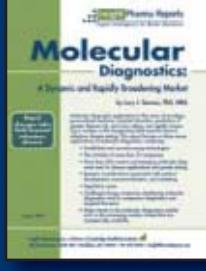
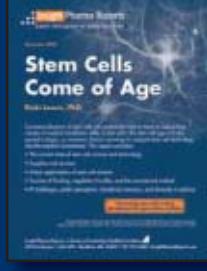
Timely • Insightful • Concise



Advance your knowledge with **Insight Pharma Reports'** collection of published reports that provide timely, insightful, and concise analysis on specific topics for your drug development needs. The reports help professionals meet the need to stay abreast of the latest advances in pharmaceutical R&D, their potential applications and business impacts, and their current and future position in the marketplace.

To order your copy(ies) of these new titles or other titles, call 781-972-5444, email rlaraia@healthtech.com, or visit InsightPharmaReports.com.

An Essential Collection of Timely & Concise Reports Analyzing the Technologies, Markets, and Strategic Issues Driving R&D Productivity



Order your copy(ies) today!

Visit InsightPharmaReports.com for other available reports.

Corporate License and Corporate Subscriptions are available. Contact David Cunningham at 781-972-5472 or Cunningham@healthtech.com for special program packages.

"CHI's Insight Pharma Reports are cost-effective and a reliable source of information about the markets we serve. The reports allow us to keep on top of developments in this rapidly moving field."

-Donald N. Halbert, Ph.D., EVP of R&D, Iconix Bioscience

Published by



Insight Pharma Reports, a division of Cambridge Healthtech Institute
250 First Ave, Suite 300, Needham, MA 02494
T: 781-972-5400 or Toll-free in the U.S. 888-999-6288 • F: 781-972-5425

Clinical Research [IN CONVERSATION]

Bleicher

(CONTINUED FROM PAGE 40)

others, and some applications may lag. It's my belief that, for whatever reason, [e.g.] academic clinical trials, specialty clinical trials, some very small clinical trials from companies trying to preserve their funds of whatever, there are going to be some clinical trials that will always be done on paper.

Phase Forward went public in 2004.

What were your key acquisitions along the way?

Lincoln Technology has a very important place in the safety arena and in the standards arena as well. It brought a lot of expertise and experience that has gathered over the years in that company. Safety, pharmacovigilance, and the ability to identify safety signals, both in clinical trials and reported data, was something that was very important to us, especially with the acquisition (before going public) of ClinSoft and the ClinTrace product. So [we've] been able to bring together the safety aspects of the ClinTrace product with Lincoln Technology's expertise and products in pharmacovigilance and in clinical trial safety, and it's made a very powerful suite, that combined with our technologies in clinical trial data management, makes an even more compelling argument.

Green Mountain Logic added a workflow aspect to phase I clinical trials, which every phase I unit needs to manage the labs and activities in a phase I clinical unit. That's very important, and I think will grow. Thirdly, the Clarix acquisition, which is in both the typical IVR systems and interactive web response systems, allows us to get randomization, more clinical trial and supply management, allows us to basically complete the loop with adaptive clinical trials. We now have everything that is necessary to do adaptive clinical trials and can offer that together to our customers in an integrated form, ready to go.

Where do you see Phase Forward's future opportunities?

...I think the company has been quite public about its interest to go to technolo-

gies and technology service providers for clinical development for post-launch safety. The idea of an end-to-end solution, technology-enabled services, actually is where the company should be focusing. There are still a number of opportunities that are available to help the company continue to consolidate and provide a one-stop shop integrated solution.

Can Phase Forward and other companies make the cost of bringing a drug to market more affordable?

Absolutely, that's been my belief all along. You're always dealing with time, quality, and cost, that triangle. Phase Forward has been able to reduce all of those, to make that triangle smaller. Each company, each clinical trial has its own interest—sometimes time is of the essence, sometimes quality, sometimes cost. The suite of techs has made it possible for companies to choose or focus on all three of them. As pharmaceutical companies get better, using that data they've collected, at data mining, looking for connections, there will be even more opportunity... That being said, the concept of ROI in EDC in general, it's almost past that point... Now people are focused on, what's the best way for my company to make this an enterprise-wide solution.

What can you tell us about your new health care venture?

Not a whole lot! We are still in stealth mode, as it's called. But I learned a lot, took a lot of principles from what we did

in the pharmaceutical industry. The next challenge is the health care industry... I've joined a company that has a really dramatic advantage, both in skills and in concept, in health care informatics... I joined as CMO, it gives me an opportunity to use a lot of things, a lot of skills in being an entrepreneur, it allows me to bring some of those principles I learned in developing Phase Forward to bear in the early days of designing and formulating a product that hopefully will change some aspects of health care the same way that EDC was enabling, and Phase Forward, was enabling for the pharmaceutical industry.

Do you think the new Obama administration will be a partner?

I do. The new Obama administration is anxious to make changes in the health care system. The current economic climate may make it difficult to make the kind of sweeping changes people are looking for, but the one thing Republicans and Democrats shared in common in their platform as they were running was electronic health records (EHRs)—connectivity through electronic data management and actually driving quality and safety of patient care through electronic data management. So I believe, if I'm reading the tea leaves right, the Obama administration is planning on making one of their infrastructure investments in high technology, in enabling EHRs and interconnectivity through electronic data. If that turns out to be correct... you heard it here first! ●

MacGarvey

(CONTINUED FROM PAGE 42)

Quantitate's statistical heritage presumably helps in that regard.

MacGarvey says Quantitate has been re-evaluating its strategic plan, but "we haven't seen any sort of downturn in the number of studies coming through. That may be serendipity, I don't know. It's becoming more competitive. Off-shoring is picking up. I've definitely seen this commoditization of some of the tasks in the biometrics area." A more pressing concern is the biotech funding affecting the pipe-

line. "That issue will hit us in a couple of years," he warns.

"We're seeing a lot more interest in protocol design. When you throw EDC into the mix, you're getting very accurate data." Growth, he says, will be bolstered by FDA pressures on trial sponsors to provide more in-depth and complex data. "The market grew by 15% last year and this growth looks set to continue," he says. "With our ability to deliver services across three continents and our flexibility to adapt to any technology environment, I believe that Quantitate is in a very positive competitive position." ●

Increase the Pace of R&D Innovation with Google

Available Now!

The R&D innovation cycle is at the heart of pharmaceutical and life sciences organizations. Yet, for most life sciences organizations, the product development process takes too long; according to the Pharmaceutical Research Manufacturers of America (PhRMA), the average drug development process can take 12 to 15 years, leaving only 5 to 8 years of patent protection. Therefore, increasing the pace of innovation has become a high priority for many organizations.

To tackle this problem, life sciences organizations are increasingly looking to IT and better tools to improve researcher productivity. In fact, according to AMR Research, 28% of IT and business managers at 61 major life sciences companies cite improving the utilization, sharing, and analysis of data across disconnected operations as a top business priority. And these organizations are increasingly turning to Google.

As the leader in search technology, Google has developed universal search for enterprises, which provides the ability to search all enterprise content – including intranets, file shares, databases, real-time business data, and content management systems – through one simple search box.

Viewers will learn about:

1. Google's tools to increase researcher productivity
2. Customer case studies of how life sciences companies have benefited from Google tools

Bio-IT World OnDemand Webcast

Featured Speakers:



David Kim
Product Marketing
Google Enterprise



Dr. Kevin Davies
Editor-in-Chief
Bio-IT World

Google™

www.Bio-ITWorld.com

Educational Opportunities

This section provides a variety of educational events in the life sciences industry that will help you with your business and professional needs. To list an educational event (print & online), contact Sales_chmg@chimediagroup.com. To preview a more in-depth listing of educational offerings, please visit www.bio-itworld.com

how to address them in a tight credit environment. From health care to energy and sustainable agriculture, biotech innovations play a major role in solving some of the world's biggest problems. Visit <http://convention.bio.org>

Featured Events

Med-e-Tel – The International eHealth, Telemedicine and Health ICT Forum, April 1–3, 2009, Luxembourg

The 7th annual Med-e-Tel focuses on a wide range of Telemedicine and eHealth experiences, business cases and research results, and attracts healthcare providers, industry representatives, researchers, and government officials from 50 countries around the world. Visit www.medetel.eu

BIO-IT Coalition's 8th Annual Conference: Will Biotechnology/IT Save the World? The Revolutionary P4 Medicine Approach to Medicine, April 22, 2009, Washington, DC

Keynoter Dr. Leroy Hood will address transformative technologies taking medicine from its currently reactive state to a personalized, predictive, preventive and participatory (P4) mode. Visit www.bioitcoalition.org/events.annual09.html

The Post-Approval Summit, April 28–29, 2009, Boston, MA

Held annually at Harvard Medical, the Post-Approval Summit focuses on strategies and best practices for demonstrating safety, quality, effectiveness, and value of healthcare products and services through patient registries, phase IV studies, risk management programs, and quality initiatives. Highlights will include: Post-Approval: The New Playing Field; Managing Safety from Surveillance to REMS; Measuring Effectiveness: Real-World Data and Comparative Effectiveness; and Practical Issues with Real-World Studies. Visit www.postapproval.org

2009 BIO International Convention, May 18–21, 2009, Atlanta, GA

The 2009 BIO International Convention covers the demands of running a biotech business in today's global economy. More than 60 countries, 20,000 life science professionals, 2,200 exhibitors and several investors will come together in Atlanta to discuss the most critical issues facing the industry and to determine

DIA 45th Annual Meeting, June 21–25, 2009, San Diego, CA Better Medicines: Improving Safety with Every Step

No matter what your area of expertise, or interest area, this meeting has been designed specifically with you in mind. Network with more than 8,500 colleagues from several geographic locations, professional interest area and level of responsibility; Listen to more than 1,100 speakers, including representatives from FDA, EMEA, and other global regulatory agencies; Choose from more than 350 sessions, 34 tutorials, and 3 mega-tracks; and much more. Visit www.diahome.org

CHI Events

For more information on these conferences and other CHI events planned, please visit www.healthtech.com



Drug Discovery Chemistry

April 6–8, 2009 • San Diego, CA

Barnett Educational Services

Visit www.barnettinternational.com for detailed information on Barnett live seminars, interactive web seminars, on-site training programs, and customized eLearning development services.

Live Barnett Seminars

Working with CROs

April 9–10, 2009 • San Francisco, CA

Pharmacovigilance in Europe: An Introduction for US Personnel

April 9–10, 2009 • San Francisco, CA

International Drug Approval Regulatory Procedures

April 27–28, 2009 • San Francisco, CA

Statistical Concepts for Non-Statisticians

April 29–30, 2009 • San Francisco, CA

Standard Operating Procedures (SOPs) in Clinical Research

May 11, 2009 • Chicago, IL

Conducting Clinical Trials Under ICH GCP

May 14–15, 2009 • Philadelphia, PA

Patient Registry Programs

May 14–15, 2009 • Philadelphia, PA

Drug Development & FDA Regulations

May 18–19, 2009 • Boston, MA

Mastering Cost Management for Global Clinical Trials

May 18–19, 2009 • San Diego, CA

Conducting Clinical Trials in Emerging Regions

June 4–5, 2009 • Boston, MA

Pharmacokinetics: A Comprehensive Overview of Principles & Applications

June 8–9, 2009 • Philadelphia, PA

Patient Recruitment & Retention: Successful Planning & Management

June 8–9, 2009 • Philadelphia, PA

Interactive Web Seminars

Operational Modeling & Simulation in Clinical Trials

April 7, 2009

Electronic Medical Records & Source Document Verification

April 17, 2009

Comparing FDA & Health Canada Regulations: Using an ICH GCP Framework

May 6, 2009

Examine Approaches for Investigator Initiated Trials

May 13, 2009

Update on Medical Device Postmarketing Vigilance Reporting

May 19, 2009

Examining the Impact of the eCTD on the Regulatory Submission Process

May 20, 2009

Approaches to Address Challenges in Vendor Management

June 2, 2009

Winning Strategies for Clinical Recruitment Media

June 9, 2009

The Pharmaceutical & Medical Device Industries Today: Comparing & Contrasting

June 16, 2009

**Recombinant Antibodies:
From Concept to Clinic**
April 8–9, 2009 • Boston, MA



PEGS – Essential Protein Engineering Summit

April 6–10, 2009 • Boston, MA

Monoclonal Antibodies

April 9–10, 2009 • Boston, MA

Advanced Biofuels Summit

April 20–21, 2009 • Washington, DC

Alliance Management Congress

April 22–23, 2009 • Boston, MA

Tools & Metrics Congress

April 23–24, 2009 • Boston, MA



Bio-IT World Conference & Expo

April 27–29, 2009 • Boston, MA

Patient Recruitment in Clinical Trials

May 11–12, 2009 • Boston, MA



Biomarker World Congress

May 27–29, 2009 • Philadelphia, PA

Protein Kinase Targets

June 1–3, 2009 • Cambridge, MA

Structure-Based Drug Design

June 4–5, 2009 • Cambridge, MA

Genomic Tools & Technologies

June 8–10, 2009 • San Francisco, CA

Translating Genomic Knowledge

June 9–10, 2009 • San Francisco

The Expanding Impact of New Animal Models in Drug Safety

June 9–10, 2009 • Philadelphia, PA



World Pharmaceutical Congress

June 9–11, 2009 • Philadelphia, PA

Risk Management & Drug Safety

June 10–11, 2009 • Philadelphia, PA

Collaborative Innovation in Biomedicine

June 22–23, 2009 • Washington, DC

White Papers, Webcasts, and Podcasts

Keep abreast of current industry trends and developments. Browse our extensive list of free Life Science white papers, webcasts, and podcasts on www.bio-itworld.com. Interested in learning about developing a multi-media solution to generate leads? Please email Sales_chmg@chimediagroup.com

White Papers

Clinical Information Fusion for R&D

Sponsored by SAIC

The white paper covers how Pharma R&D's information silos can be broken to provide on-demand, up-to-date information supporting immediate business needs, utilizing existing technology investments, and reducing future implementation and maintenance costs.



Adobe® LifeCycle® ES—Deploying Adobe Technology to Automate Electronic Submissions

Sponsored by Adobe

The white paper focuses on how to reduce the cost of bringing products to the market; improve document collaboration across your organization; satisfy global regulatory requirements, stockholder expectations, and customer demands; improve data capture, information assurance, document output, process management, and content services; and gain a competitive advantage and achieve immediate ROI.

Interoperability and Architecture for the Life Sciences Industry – What it will take to gain industry consensus on interchange standards, and what organizations can do today to optimize their own information Architectures

Sponsored by SAS

Information technology is finally up to the task of cost-effective clinical research. The white paper explores the need for a consistent industry architecture that allows life sciences firms to connect their enterprise and benefit from unified data integration, process consistency, and rapid communication of meaningful drug findings.

Digital End-to-End for Pharma

Sponsored by SAIC

How can Pharma effectively harness the immense amounts of data from an ever-growing number of internal and external sources? The white paper covers options where Pharma can harness the power of data to pursue novel products that are profitable, safe, and have proven value to one or more patient populations.

Webcasts

Electronic Submissions – Adobe's Prescription for Insuring Success

Sponsored by Adobe

In the time-consuming and costly battle to gain market share, faster time to market equals a competitive advantage for life sciences companies. Adobe® LifeCycle® ES (Enterprise Suite) solutions help simplify the submission process, enabling life sciences companies to achieve earlier filing dates, greater accuracy and quality of data, and ultimately faster time to market.



Increase the Pace of R&D Innovation with Google

Sponsored by Google

Life sciences organizations are increasingly looking to IT and better tools to improve researcher productivity and shorten the product development process. As the leader in search technology, Google has developed universal search for enterprises, which provides the ability to search all enterprise content—including Intranets, file shares, databases, real-time business data, and content management systems—through one simple search box.

Podcasts

Next Generation Sequencing Data Management

Co-Sponsored by Illumina and GenoLogics

Targeting genomics labs with increasing volumes of data and complexity of experiments, this end-to-end solution helps manage their entire process from samples to results. Learn how the fully integrated Genome Analyzer-Geneus solution can help maximize your next generation sequencing investment.

Rising Clinical Trial Delays and Costs – Addressing the Cause, Not the Symptoms

Sponsored by Medidata Solutions

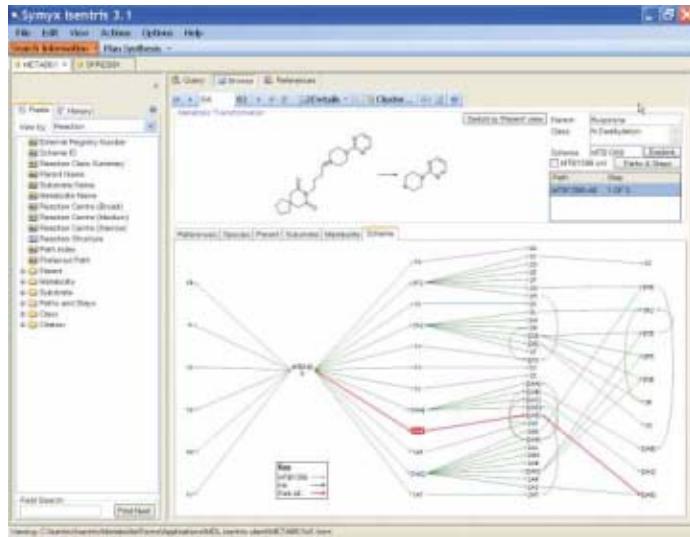
Protocol complexity is taking a toll on clinical study speed and efficiency: increasingly complicated and ambitious protocols are not only burdening sites and study volunteers, but also prolonging trials and increasing expenses. View the podcast to learn how sponsors are responding to this critical issue.

Manage Your Biorepositories

Sponsored by GenoLogics

Learn about solutions for managing your bio-specimens and biorepositories such as: What is the single largest challenge facing organizations in terms of accelerating research; How is data collection being managed across disparate organizations and sites; How will data be managed in terms of complying with regulations (e.g., HIPAA).

New Products



New Databases

Symyx has announced the release of two new databases. SPRESI is a structure and reaction database from InfoChem GmbH, available through Isentris. SPRESI extends the existing set of reactions available from Symyx by providing access to an additional 6 million structures, 3.8 million reactions, and 28 million factual data entries extracted from 636,000 references and 164,000 patents. The latest version of the Symyx Metabolite database is also now available through Isentris.

Metabolite is the largest and most comprehensive collection of drug metabolism data in the world, and users of the latest release of Metabolite can access all Symyx molecular property, chemical sourcing, synthetic methodology, bioactivity, and metabolism/toxicology databases.

Product: SPRESI and Metabolite databases
Company: Symyx
Available: Now
For more information: www.symyx.com

Sensitive Spectrometer

Waters Corporation has introduced its Waters Xevo quadrupole time-of-flight (QToF) mass spectrometer (MS), an exact mass MS/MS benchtop instrument that is the most sensitive benchtop QToF system ever. Paired with Waters ACQUITY UltraPerformance Liquid Chromatography (UPLC), the Xevo QToF MS is the only commercially-available MS system to give scientists one-of-a-kind UPLC/MSE performance—a novel, patented method of data acquisition that captures “all of the data all of the time” at unparalleled speeds to maximize the amount of information collected from minimal sample amounts.

Product: Xevo Quadrupole Time-of-Flight Mass Spectrometer
Company: Waters
Available: Now
For more information: www.waters.com/xevo

Genotyping Suite

The TaqMan OpenArray Genotyping System expands the potential uses of TaqMan technology across a wide range of genotyping applications, providing an extremely fast, high-sample-throughput validation and screening tool for researchers in agricultural, pharmaceutical, and other commercial industries, as well as academic institutions. The system includes an instrument platform, reagents, and consumables. The main components are the TaqMan OpenArray Genotyping Instrument Platform and TaqMan OpenArray Genotyping Plates, which are pre-loaded with pre-designed or custom genotyping assays. Researchers using this integrated platform can experience an end-to-end genotyping workflow of less than eight hours, enabling them to analyze thousands to tens of thousands of samples in days, in contrast to weeks or months, on alternative genotyping platforms.

Product: TaqMan OpenArray Genotyping System
Company: Applied Biosystems
Available: Now
For more information: info.appliedbiosystems.com/taqmanopenarray

The Supercomputer is Personal



NVIDIA and its worldwide partners today announced the availability of the GPU-based Tesla Personal Supercomputer, which delivers the equivalent computing power of a cluster, at 1/100th of the price and in a form factor of a standard desktop workstation. Priced like a conventional PC workstation, yet delivering 250 times the processing power, researchers now have the horsepower to perform complex, data-intensive computations right at their desk, processing more data faster and cutting time to discovery. At the core of the GPU-based Tesla Personal Supercomputer is the Tesla C1060 GPU Computing Processor which is based on the NVIDIA CUDA parallel computing architecture. CUDA enables developers and researchers to harness the massively parallel computational power of Tesla through industry standard C.

Product: Tesla Personal Supercomputer
Company: NVIDIA
Available: Now
For more information: www.nvidia.com



Bio-IT World Best Practices Awards Dinner and Ceremony

A Special Invitation



The **Bio-IT World Best Practices Awards** Networking Reception and Dinner recognizes teams for their outstanding innovations and excellence in the use of technologies, practices and novel business strategies that will advance drug discovery, development, biomedical research, and clinical trials.

Please join us for this special event and share in the excitement as prestigious awards are presented acknowledging innovative activities accomplished in the life science community. Dinner attendees will have the opportunity to network with life science colleagues and receive the 2009 Best Practices Awards compendium showcasing the sixty-nine entries submitted.

"[The Best Practices Awards] is an indication of how we are all working together in this era and that is a wonderful thing."

-Dr. Francis Collins
Natl Hum Gen Res Inst, NIH

Date: Tuesday, April 28, 2009

Time: 7:00 – 7:45 p.m. – Networking Cocktail Reception
7:45 – 10:00 p.m. – Dinner and Award Presentations

Venue: World Trade Center
200 Seaport Boulevard
Boston, MA

Room: Cityview Ballroom

Guest Speaker: Michael Cariaso, BioTeam
Technology Without Borders: Teaching Programming
and Internet Skills in Rural South East Asia



Book Your Dinner Reservations Today!

For dinner and reservation details,
visit <http://www.bio-itworldexpo.com/awards.asp>,
email jchambers@healthtech.com, or call 781-972-5400.

Awards Dinner Sponsors



The Russell Transcript



CoSBi Models

JOHN RUSSELL

Using computers to model living systems and to make useful predictions about their behavior is a holy grail pursued by many; yet there are far fewer seeking to go in the reverse direction and infer principles from living systems to improve computer science. Corrado Priami, president and CEO of the [Centre for Computational and Systems Biology](#) (CoSBi), a partnership between Microsoft, the Italian Government and [University of Trento](#), is hoping to do both.

Formed just three years ago, CoSBi has grown steadily (see, “[The CoSBi Show](#),” *Bio•IT World*, Dec 2006). Today it has about 25 researchers and has already developed a number of prototype tools freely available at www.cosbi.eu: BetaWB, Cytosim, Kinfer, Snazer, and Redi.

More recently, Priami has written a short paper on “algorithmic systems biology” that will be published in *Communications of ACM* (Association for Computing Machinery) this June, in which he sketches his ideas on how computer science can be brought to bear on biology and vice versa. The heart of what computer science can learn from biology, he says, is parallelism and robustness. “Computers crash too easily, they’re not naturally tolerant, while living systems are robust and adapt well to environment. Maybe we can make better software using principles from living systems,” he says.

Priami says that the convergence between computing and systems biology provides “a valuable opportunity that can fuel the discovery of solutions to many of the current challenges in both fields, moving towards an algorithmic view of systems biology.” He envisions different levels of cross fertilization between the two areas. Living systems are much more robust than current computer systems, much more adaptive to their environment. A priority—“probably of interest to our joint venture with Microsoft”—is learning how to improve the tolerance of software development. Another lesson, he suggests, might be more energy-conserving approaches to computation.

In that paper, Priami picks up the ongoing debate over whether computer science is indeed a science, and contrasts it to mathematics with which computer science is often allied in a subordinate, supportive role. He argues computation is a distinct science, in part because its operations necessarily have

a physical reality while mathematics may be theoretical or solvable in the abstract. He further argues that living systems share much with computation since they too are information processing (i.e. computation) systems in which real events must occur to do the “computation.”

Borrowing from Biology

Broadly speaking, CoSBi is trying to develop a new programming language, syntax, and toolset to model and simulate living systems. Asked to distinguish his approach from companies such as [Entelos](#), [Genstruct](#), and [Gene Network Sciences](#), Priami says: “We are not using mathematical tools; we are using computer science tools. There are lots of professional reasons for doing that because biological systems are highly parallel; you have thousands of interactions that happen simultaneously. Mathematical modeling is mainly an equation and it is combinatorial in size so when you have to describe all those ways in which systems can touch, it is not so suitable to model large systems.”

Priami says he is trying to produce models that are “transition-based, not state based.” State-based systems, he says, “describe the differential expressions or other quantities of variables with state changes and try to condense this into relations and [from these] infer the dynamics of the systems. We’re trying to describe not the state, but the transition from one state to another, which is exactly the way computer scientists describe the behavior of distributed programs, so programs that run on different nodes of a network, exchanging messages. It is a newer way of representing the phenomena.”

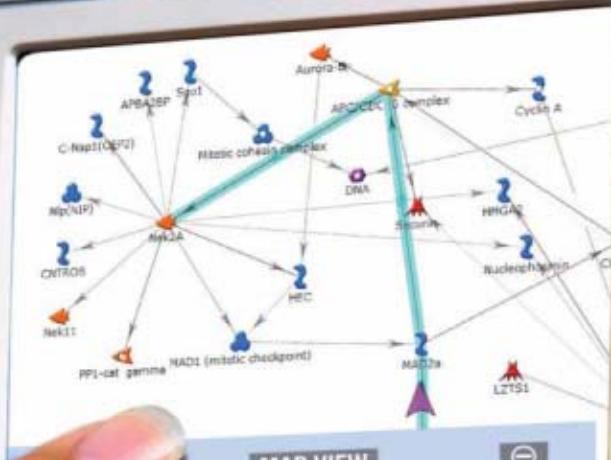
“Our approach is much more scalable. We do not have theoretical limitations to the size of the model; only technology limitations in terms of the size of the memory we can use. We aim to make models [that have], say, gene regulatory networks and metabolic networks and what people ordinarily do is look at the networks in isolation. We attempt to have the two together interacting in the same system so that we can understand how they interact together. A long term perspective is to be able to model things like immune system at a molecular level.”

Ease of use remains an issue, agrees Priami. In the next 18 months, he hopes to develop an interface to the platform that is useable by biologists. He also hopes to validate some *in silico* predictions in the lab and demonstrate effective crosstalk between large systems that demonstrate the scalability of CoSBi’s approach.

It should be fascinating to watch not only the systems biology tools developed by CoSBi but also to monitor how successful it is in taking ideas and techniques from living systems and incorporating them into computer science.



Corrado Priami



Your GPS in Pathways Analysis

MetaCore™

Gene Expression and Pathway Analysis Software

GeneGo's unique canonical pathway algorithm gets you to your destination quickly using well established biological highways. MetaCore's integrated data mining platform puts you in the driver's seat with the tools you need for functional analysis of gene expression, metabolomic, proteomics, siRNA, SNP, HCS and other types of experimental data in the context of manually curated pathways. MetaCore is easy to use and includes free on-line training to get you on the road and up to speed.

Test drive MetaCore™ today with a FREE trial – trials@genego.com



500 Renaissance Drive, Suite 106
St. Joseph, Michigan 49085 USA
(888) 592-3124 , (858) 756-7996
+44 7525 064993 United Kingdom
www.genego.com
sales@geneego.com

Go beyond the everyday. **EVERY DAY.**



Symyx Notebook

The Freedom to Experiment.

Whether it's on the trail or in the lab, you want the freedom to take new approaches, routes, and paths to your goals. That's why there's Symyx Notebook. It's the only electronic laboratory notebook that can be deployed across the enterprise in multiple scientific disciplines. With Symyx Notebook, research teams share a single application to document, work, collaborate, and speed the experimentation workflow.

Symyx Notebook streamlines the capture of all experimental information and intellectual ideas. Everyday tasks such as data capture and note taking are optimized and automated. All of which gives you the time and freedom you need to experiment—and get back to doing science.

*To learn more, visit
www.symyx.com/notebook6*

Symposium 2009

Register Now and Save at
www.symyxsymposium.com



© 2008 Symyx is a registered trademark of Symyx Technologies, Inc. All rights reserved.