Bio·IT World
Technologies Driving Discovery, Development, and Clinical Trials

Welcome to your digital edition of *Bio-IT W*

We hope you Bio-IT World community, y yer's Guides, Events, and C COM for daily delivery of fre dlines on your desktop (or yo

In addition, y covering such targeted top s, Regulatory Compliance, mplete list of enewsletter ti sign-up.html

And as always nd us an email with DIGITAL

>> CLICK
>> CLICK H

Bio-ITWorld reports the following digital circulation figures as verified by the Qmags Self-Audit procedures:

November 2007 Issue	
Total qualified digital subscribers:	6343
Biopharma/Chemicals	2015
Biotech/Bioinformatics	1805
Acad/Gov, Clinical Research	1746
Hosp Lab/Healthcare, Other	1077
Oct 2007 Issue	
Total emails sent:	7202
Undeliverable:	73
Total delivered:	7129
Removed:	27

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
- Zoom In: top will zoom in to the top of page & bottom will zoom in to the bottom of page
- Zoom out: takes you from single to double page view
- Front Cover: takes you to the Cover Page of magazine
- Search Issue: allows full search by keyword or phase (can also search archives!)

WE CAN. Yoh has been delivering high-impact talent to America's leading companies for the past seven decades – longer than any of our competitors. These customers count on our experience, time-honored processes, responsiveness and market intelligence. And our delivery of talent that is 24.7% better than the competition*.

Focus on your job. And your future.

YohClinical
Know greater talent®
yohclinical.com

Cambridge Healthtech Institute

www.bio-itworld.com

Bio·IT World

Indispensable Technologies Driving Discovery, Development, and Clinical Trials

NOVEMBER 2007 • VOL. 6, NO. 9

HIGH-CONTENT ANALYSIS

Software Solutions for Cellular Screening 26

+

CLOUD COMPUTING AND LIFE SCIENCES 18

RUDY POTENZONE: MANAGING THE BIOIT ALLIANCE 31

TRAINING DAY: SYSTEMS IN SEATTLE 20

OPEN-SOURCE INTELLIGENCE FOR PHARMA 16

Connect the dots

PathwayArchitect® Software from Stratagene—make finding relationships simple

TRY IT FREE TODAY!

Download a trial version today at www.stratagene.com/softwaresolutions

u.s. and canada 800-424-5444 x3
europe 00800-7000-7000

www.stratagene.com

Our next-generation PathwayArchitect® Software is designed for the life scientist who needs to obtain facts about a disease-relevant biological pathway as well as understand the statistically relevant relationships in their system and integrate complex numerical data. The guided workflow within the software takes you through the steps of searching for your pathway of interest. You can also build a pathway based on a list of genes and utilize the software to integrate and analyze your numerical data.

- **Pathway database expands target discovery**
- **Uncover new biological relationships**
- **Visualize your expression data as pathways**

Visit www.stratagene.com/softwaresolutions for complete details about our PathwayArchitect® Software and information about Academic, multiple year, and multiple user licensing options—and connect the dots.

STRATAGENE
An Agilent Technologies Company

Strand Life Sciences

Count on Us.

Over 1500 customers worldwide already count on Thermo Scientific informatics solutions to manage one of their most important assets – laboratory data. These industry leaders have turned to Thermo Scientific LIMS and CDS software in their efforts to make the world healthier, cleaner and safer.

If you are looking for industry experience, domain expertise, innovative technology and world class after-sale support, you can count on us to meet your needs now, and well into the future.

Find out which Thermo Scientific informatics solution is right for you. Visit **www.thermo.com/solutionfinder**.

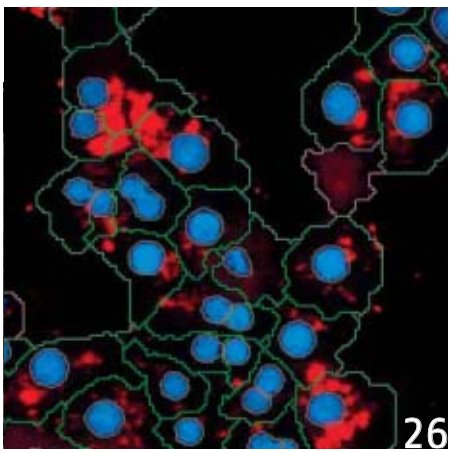
+1 866-463-6522 • marketing.informatics@thermofisher.com

We have a solution for you, no matter what industry you are in:
pharmaceuticals, food and beverage, industrial chemicals, petrochemicals, metals and mining, water and wastewater and many many more.

Part of Thermo Fisher Scientific

Thermo SCIENTIFIC

Contents [11·07]



Cover Story

High Content Analysis

As HCA matures, developers walk a thin line between power and ease of use. 26

Up Front

Computational Biology

Computational Development

IT/Workflow

Clinical Research

In Every Issue

Alternative Thinking About Server Rooms:

ADDITION THROUGH SUBTRACTION.

Take away the jungle of cables. Take away the so-called "normal" energy requirements and maintenance. What do you have? Introducing the HP BladeSystem c3000. All the technology of our larger BladeSystem in an efficient, compact, affordable package.

Technology for better business outcomes.

Powered by the Quad-Core Intel® Xeon® Processor[1]

See how less is more. Visit hp.com/go/nocompromise21
1-888-335-6443

1. Intel, the Intel logo, Xeon, and Xeon Inside are trademarks or registered trademarks of Intel Corporation in the U.S. and other countries.

First Base

Can Academics Save Pharma?

KEVIN DAVIES

Two big pharmas — Pfizer and GlaxoSmithKline (GSK) — made critical executive hires last month. After a yearlong search, GSK ended up plucking company insider Andrew Witty, a 40-something former head of Pharmaceuticals Europe, to succeed retiring CEO Jean-Pierre Garnier. GSK's pipeline is considered strong, but the company is still trying to recover from the recent Avandia controversy.

Meanwhile, Pfizer, facing the looming expiration of $12-billion-per-year Lipitor patent, is urgently looking to reinvigorate its pipeline (which consumes $7 billion annually). Thus, the news of the hiring of Martin MacKay as the company's new R&D director and former Merck executive Briggs Morrison to head up clinical development was noteworthy. But even more intriguing was the appointment of Corey Goodman to head a new Pfizer research center in the Bay Area. Goodman was the founder and CEO of Renovis, a San Francisco biopharma producing drugs for pain and stroke. Goodman bowed out of Renovis following its $152-million acquisition by Evotec AG, only to resurface at Pfizer a few days later. He will report directly to CEO Jeffrey Kindler.

It's probably fair to say Goodman isn't being hired for his Renovis record as much as his outstanding academic credentials and creativity. For two decades, Goodman's group at first Stanford, then Berkeley, led the study of axonal migration, making key discoveries in the identification of molecular cues that guide the nervous system development, albeit in fruit flies. In 1994, he co-founded Exelixis, another Bay Area biotech, with two fellow fruit fly geneticists.

Goodman's Bay Area center will bridge the gap between basic research and drug discovery, Kindler said, partnering with academic groups and incubating startups with innovative technologies. The latter sounds reminiscent of Biogen Idec's BI[3] incubator (See "Biogen Idec's Innovation Incubator," *Bio•IT World*, March 2007). Goodman said the Bay Area center would be independent, "free to establish its own distinct culture, and empowered to recruit entrepreneurial scientists." It would also leverage the assets of Pfizer Global R&D (PGRD), using the company's high-throughput screening and pharmaceutical science capabilities and collaborating with PGRD's biotherapeutics teams, before handing off new drug candidates for late-stage clinical development.

Goodman added: "While we will be focused on biotherapeutics, we will look for any innovative technology in any area that will help develop new medicines. We will be in the center of the California biotech and venture community, in the midst of some of the greatest biomedical research institutions."

> We will look for any innovative technology in any area that will help develp new medicines.
>
> **Corey Goodman**, Pfizer

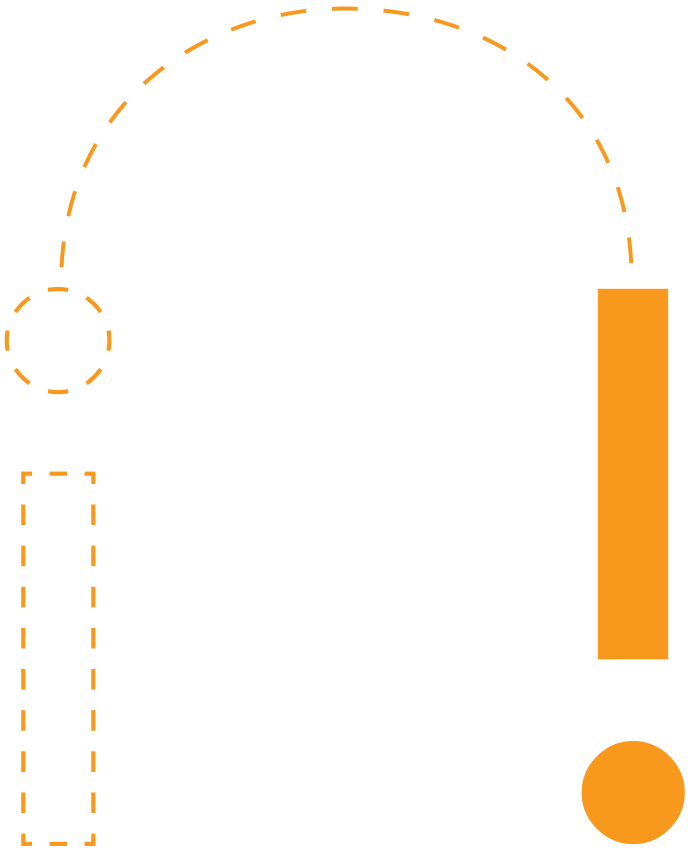
Question Time

This surprising turn of events raises many questions, not least being the time it will take for Goodman to make his mark. Pfizer owns the Rinat Neuroscience facility in South San Francisco, following the acquisition of the protein therapeutics firm in 2006. The site is a key part of Pfizer's biotherapeutics program spanning various disciplines including oncology, neurology, infectious and metabolic diseases. Another uncertainty is how Goodman's new center will dovetail with Pfizer's existing Research Technology Center (RTC) in Cambridge, Mass., though the infusion of new blood and ideas must be welcomed.

It's a fascinating gamble from Pfizer, which saw its plans to replace Lipitor with Torcetrapib crash last year. But Goodman's hiring has ample precedent: several pharmas have turned to world-class academics to rethink their approach to drug discovery. Novartis hired cardiologist Mark Fishman to head its Cambridge research hub, the Novartis Institute for Biomedical Research. Despite some high-profile losses — former oncology chief Sasha Kamb and functional genomics director Mark Boguski have both left — Fishman continues to lure leading academics to the institute. The most recent example is Mark Keating, a renowned cardiologist at Children's Hospital/Harvard Medical School, who now runs ophthalmology.

At Genentech, Richard Scheller and Marc Tessier-Lavigne head a research program that has seen great success in the biologics arena, prompting other pharmas to emulate this approach. Another success story appears to be taking shape at Merck, where former Whitehead Institute structural biologist Peter Kim is turning the research program around, even as the Vioxx saga drags on. Interestingly, all of the aforementioned recruits were investigators with the Howard Hughes Medical Institute. GSK has also been down this route before, hiring former Cambridge University chair of genetics Peter Goodfellow to head discovery research in 2000. Goodfellow led an international team of some 1500 scientists until 2006.

The scale and complexity of global drug development makes it difficult to gauge the long-term success these academic infusions are having. Pfizer has certainly picked a superb scientist for the task, but time is of the essence.



Turn sequencing on its head.

It's here. Now you can perform a range of genomic experiments that you only dreamed of before. The Illumina Genome Analyzer changes the way you approach every kind of genomic experiment. The discoveries are happening, the wait is over, and our customers are publishing their results.

Change the way you think about sequencing.
Join the growing Illumina Community.

Do more with sequencing today:

www.morethansequencing.com

illumina®

Company Index

Advertiser Index

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

VOLUME 6, NO. 9

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

Bio·IT World (ISSN 1538-5728) is published monthly except July/August and December/January 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 inquires to Bio-IT World, P.O. Box 3414, Northbrook, IL 60065 (888) 835-7302 or e-mail biw@omeda.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

EXECUTIVE EDITOR
John Russell (781) 972-1342
john_russell@bio-itworld.com

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

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

WEB EDITOR
Catherine Varmazis (781) 972-1344
catherine_varmazis@bio-itworld.com

DIRECTOR, BUSINESS DEVELOPMENT
Angela Parsons (781) 972-5467
aparsons@healthtech.com

VP SALES – CA, WESTERN US, CANADA, PACIFIC RIM
Alan El Faye (213) 300-3886
alan_elfaye@bio-itworld.com

REGIONAL SALES MANAGER – MIDWEST, SOUTH EASTERN US, EUROPE
Jennifer L. Wickett (781) 972-5482
jwickett@healthtech.com

REGIONAL SALES MANAGER – NEW ENGLAND, NORTH EASTERN US
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
Kathrene B. Tiffany (781) 972-5440
ktiffany@healthtech.com

GRAPHICS AND PRODUCTION MANAGER
Ann-Marie Handy (781) 972-5493
ahandy@healthtech.com

Contributing Editors
Michael Goldman, Karen Hopkin, Deborah Janssen, Pauline Parry, Salvatore Salamone, Tracy Smith Schmidt, Mark D. Uehling

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@bio-itworld.com
250 First Avenue, Suite 300
Needham, MA 02494

You Stand in Front of Our Instruments All Day... We Stand Behind Them...24x7

High Value, Substantially Expanded Services from Applied Biosystems

With over 25 years of experience in the development, manufacture, and service of innovative instruments, and with over 1000 highly trained, dedicated service professionals, AB Global Services is uniquely qualified to deliver a full suite of real-world lab services. From Remote Services to On-Site Application Consulting to Qualification and Professional Services and everything in between, AB Global Services is your value-added partner to help boost your productivity and maximize the return on your technology investment.

- Instrument Repair and Maintenance
- Smart Services
- Qualification Services
- Professional Services
- On-Site Application Consulting
- Training

To learn more go to http://info.appliedbiosystems.com/wwservice or contact your local Applied Biosystems sales representative.

Applied Biosystems

Renewed Purpose for Some Drugs

CHI's Drug Repositioning Summit found new tricks for old compounds.

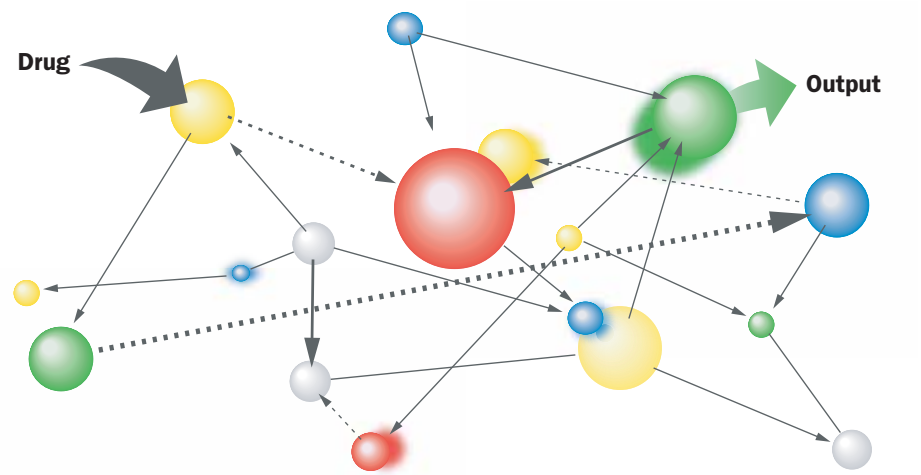
BY KEVIN DAVIES

PHILADELPHIA — "The most fruitful basis for the discovery of a new drug is to start with an old drug," the Scottish pharmacologist and Nobel laureate James Black once said. A number of speakers at CHI's second annual Drug Repositioning Summit* quoted that phrase, which begged the question that **Chris Lipinski**, the former Pfizer chemist of "Rule-of-Five" fame, duly asked: "Why, then, has drug repositioning only become common in the past four or five years?"

Given the problems facing the pharma industry, it is curious that big pharma and more opportunistic biopharma brethren only seem to have grasped the potential of repositioning in the past few years. Companies such as Aspreva, Gene Logic, KineMed, Melior Discovery, NV Organon, and Sosei in Japan have all made drug repositioning a central tenet of their business plans and more companies are destined to emerge.

In response to Lipinski's rhetorical question, **Donald Frail**, head of Pfizer's Global Indications Discovery Unit in St. Louis, offered two reasonable answers: one, because the promise of the genomics revolution was oversold. The second was emergence of new "business realities" — in other words, the most important "omics" of all — economics.

Whatever the drivers, drug repositioning and repurposing are gaining momentum. **Marcel van Duin**, executive director with NV Organon, which is in the process of being acquired by Schering-Plough, called it "a shift from target-based discovery to compound-based discovery." Organon is pursuing several repurposed drugs for new indications including contraceptives and anesthetics.



Pathway-focused discovery recognizes that a drug affects a pathway, not an isolated target.

> "We do work on failed compounds, we just don't know which ones have failed yet!"
>
> **Donald Frail**, Pfizer

Lipinski noted some 80 or more ongoing repositioning projects, including some of the most high profile drugs on the market — Tamoxifen for bipolar disorder, Gleevec for rheumatoid arthritis, and Lipitor for Alzheimer's disease and influenza. "One of the advantages of repurposing is that you've already invested so much," said the veteran chemist.

Lipinski is now a scientific advisor to Melior Discovery, a biopharma in Exton, Pa. Melior uses a multiplexed *in vivo* screening platform called theraTRACE to screen compounds across more than 30 potential therapeutic indications. Melior's co-founder and VP research, **Michael Saporito**, said the company's approach works on most therapeutic areas, although one notable exception is oncology, "because *in vivo* models are not that predictive for use in the clinic."

Saporito said the company has assessed more than 150 compounds so far. MLR-1023 — the company's lead product — is a chemical that Lipinski originally synthesized at Pfizer in the 1970s. The compound withdrawn from phase III trials. The theraTRACE screen revealed a hit in the glucose tolerance test. Further studies showed it activates the Lyn kinase signaling pathway *in vitro* and could make a good combination therapy drug for diabetes.

A Biotech Within Pfizer

Pfizer is taking an even more radical approach towards drug repurposing. Under Frail, Pfizer has assembled a dedicated team of 50 scientists in St Louis to identify new indications for compounds — both failed and those still active in the clinic.

Frail said "indications discovery" was first used at Pfizer as far back as 1999, but a dedicated division was created in 2006. "Drug repositioning," said Frail, referred to failed compounds, whereas the goal of "indications discovery" is to expand the indications for active and inactive compounds. "We do work on failed compounds," he said, "we just don't know which ones have failed yet!" Frail likened the new St. Louis center to "a biotech within Pfizer."

The new group combines wet lab ex-

(CONTINUED ON PAGE 14)

*CHI's Drug Repositioning Summit: Finding New Routes to Success. October 10-11, 2007; Philadelphia, PA.

Sequenom's Eye on the Future

CSO Charles Cantor talks about nanopore technology.

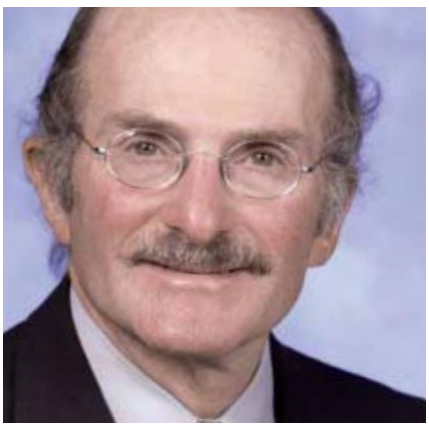
For much of the past year or two, Sequenom has been making headlines for all the wrong reasons, flirting with delisting from the NASDAQ exchange. But recently the mass spec analysis company has rebounded strongly, its share price jumping above $8. In September, news came that the company is looking to commercialize nanopore technology. According to president and CEO Harry Stylli, "Long term, we believe it has the potential to provide a commercially viable, rapid, sub-thousand dollar human genome sequencing solution."

Bio•IT World *Editor-in-Chief Kevin Davies asked Sequenom CSO Charles Cantor about the firm's recovery and future plans.*

Bio-IT World: So how is Sequenom doing these days?
Cantor: Well, we're doing much better — our performance both in terms of revenue and on the market clearly shows this. We have bounced back.

What do you attribute that to?
Well, I think we got serious, and really improved our product and made it much more reliable and broadened the range of applications. We have a lot of happy customers now, and a couple of years ago we had many unhappy customers! People have used the Sequenom platform for genotyping, for quantitative gene expression, for methylation measurements. We're about to launch a product for bacterial and viral identification. So it's not just a SNP platform now.

Where does the interest in nanopore sequencing come from? Is this just one of several long-range technologies you're exploring?
It's not the only future technology we're looking at, but it's one I'm very attracted by. Technology is not static, OK? The mass spec is great for studies of a relatively limited number of markers on large numbers of samples. We think our sweet spot is 100 to 1,000 loci — whether they be SNPs, expression, methylation, it doesn't matter — anywhere from hundreds of samples to tens of thousands of samples. That's what our customers use our platform for. But the mass spec is not competitive today where you have to look at larger numbers of markers simultaneously.



And of course, some array companies have 1 million SNPs on a single array...
Exactly. So the way I'm viewing the nanopores from the 60,000 feet level is, I think I can use them as virtual arrays. So that I would be able to get the same density of information out of them that you can with the arrays, but I have the tremendous advantage that everything is in homogenous solution. Sometimes nucleic acids work better in homogeneous solution; they don't like to be on silicon surfaces.

So you're not necessarily looking at this as a sequencing technology per se?
That's correct, I'm looking at it as a broad enabling platform for all kinds of things. Just as the mass spec is a broad enabling platform for all kinds of things. We do sequencing on the mass spec too — that's how we do methylation. To me it was intellectually a very good fit. It has homogeneous input, no surfaces — and the major challenge on the output side is very fast data processing — digital signal processing, which we're good at. We analyze our mass spec data on the fly, and we think the nanopore data's going to have to be analyzed on the fly for optimum use.

Who have you partnered with and what are the next steps?
We licensed the technology from Boston University and Harvard. Our academic partner is Amit Meller, who developed this approach while at Harvard, he is now in the engineering school at BU. Meller was a postdoctoral fellow with Dan Branton (Harvard). Branton was the first person to get the nanopores to work, to visualize DNA molecules. The focus of that lab has largely been on electrical detection. Amit is an optical physicist, who transitioned that platform from electrical detection to optical detection.

How will this be developed?
Our intent is to move forward aggressively, by collaborating with people externally and in house. The long-range plan of course is to develop a mature platform for DNA sequencing. My guess at the moment is that other applications will come quicker... I view this as a 3rd generation system. I think it's a broad enabling platform, and I think it may be a faster route to market for genotyping and gene expression than it is for sequencing. The example I'll give you is the arrays are very good for most things, but very few people are using arrays to try to do sequencing. The way I view these nanopores is virtual arrays, so it's the same kind of argument.

Do you think nanopores could eventually become a viable 3rd-generation sequencing platform?
Oh, I think they're potentially revolutionary for sequencing, and we plan to develop them for sequencing, don't get me wrong! But if I ask how long will it take to have a commercial sequencing platform that I can sell as opposed to how long will it take to have a nanopore that's not quite as demanding as whole-genome sequencing, then the answer is, I bet the first products will be more limited than nanopore sequencing. ●

Technology Leadership in High Content Analysis (HCA) with Frost & Sullivan and Definiens

Available Now!

HCA data volumes can be astounding, demanding robust high throughput solutions that can deal with data variability and biological complexity. Because most HCA data is created as images, technologies that support and advance high content analysis through image analysis are a focal topic for discussion.

Join this webcast to hear:

- Analysis from Frost & Sullivan concerning HCA technology leadership
- Approaches for dealing with data complexity in a high-throughput environment
- How Definiens context based image analysis is supporting and advancing HCA at AstraZeneca, the University of Oxford, and MIT

Participants will:

- Get the latest analyst perspectives on high content analysis from Frost & Sullivan
- Learn about industry leading image analysis software for HCA offered by Definiens
- Leverage insights from HCA work being done at AstraZeneca, the University of Oxford, and MIT

Bio-IT World
OnDemand Webcast

www.bio-itworld.com/webcasts/definiens

Featured Speakers:

Dr. Colin Blackmore
Application Scientist
Definiens

Jonathan Witonsky
Industry Manager & Analyst
Drug Discovery Technologies and Clinical Diagnostics
Frost & Sullivan

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

FROST & SULLIVAN

DEFINIENS®
Understanding Images

Repositioning

(CONTINUED FROM PAGE 10)

perimentation and *in vivo* models with a heavy dose of informatics and computational biology. Frail noted the early value of a text-mining tool called PharmaMatrix, as well as other tools for expression analysis, systems maps, and clinical informatics. He also praised Pfizer's first-generation information system for compound data, which was recently deployed (See p. 34).

But Frail also issued a warning: "In our enthusiasm [for repositioning], let's not forget it is all about Phase 2 success. Is there any reason to believe that a repositioned candidate will have less attrition due to lack of efficacy in Phase 2?"

Integrative Pharmacology

Tom Barnes, senior VP discovery at Gene Logic, said his firm is pursing "integrative pharmacology" with a host of partners including Pfizer, Roche, Organon, Lilly, Lundbeck, and Merck Serono. Gene Logic uses a variety of technology platforms, including *in vivo* imaging, multiplex bioanalytics, molecular and cellular pharmacology, and *in silico* biology. Its BioExpress database, for example, contains expression data on more than 20,000 samples and 400 disease states. A former Millennium Pharmaceuticals compound, GL1001, originally developed as an ACE2 inhibitor, is being developed for pancreatitis.

Not surprisingly, a panel discussion on the topic of IP of drug repositioning featuring **Richard Smith**, a partner with Edwards Angell Palmer & Dodge, generated intense interest. Smith said that composition-of-matter patents on drugs are the strongest. On the other hand, methods-of-use patents are relatively weak, subject to off-label uses that erode protection. New court and patent office rules make it much easier to find "obviousness" and restrict lengthy presentations.

Judging from the stream of questions for Smith, there is new purpose among those pursuing repositioning strategies. ●

Bio•IT World Expo Announces 2008 Keynote Speakers

Avey, Boger, and Reynders head sixth annual Expo.

BY BIO-IT WORLD STAFF

Bio•IT World has announced the three keynote speakers for its sixth annual Conference & Expo, to be held in Boston next spring (April 28-30).

Linda Avey is the co-founder of 23andMe, the Bay Area consumer genomics start-up. She previously worked in business development for Affymetrix and Perlegen Sciences. Prior to that, Avey had stints at Spotfire and Applied Biosystems. She founded 23andMe with Anne Wojcicki, wife of Google co-founder Sergey Brin (Google is also an investor in 23andMe).

Wednesday morning, April 30, 2008.

Joshua Boger, Ph.D., is founder, president, and CEO of Vertex Pharmaceuticals. Boger has served as CEO since 1992, and has also served as chairman and president, and chief scientific officer. Prior to founding Vertex in 1989, Boger was senior director of basic chemistry at Merck Sharp & Dohme. His early years at Vertex were featured in Barry Werth's classic biotechnology book, *The Billion Dollar Molecule*.

Tuesday morning, April 29, 2008.

John Reynders, Ph.D., is the newly appointed CIO for Johnson & Johnson's Life Sciences Division. He was previously the information officer for Lilly Research Laboratories — discovery and development informatics. Reynders joined Lilly from Celera, where he was VP for informatics and information systems, where he was responsible for all supercomputing capabilities, discovery software engineering, and enterprise system infrastructure.

Monday afternoon, April 28, 2008. ●

Briefs

GENSENSE FOR GENOMICS

The University of British Columbia's James Hogg iCAPTURE Center for Cardiovascular and Pulmonary Research has chosen InforSense's GenSense workflow-based analytics platform to support genetic research programs. The center, based at St. Paul's Hospital in Vancouver, hopes GenSense will aid in analyzing large data sets generated by genotyping.

HOPEFUL OUTLOOK

Thermo Fisher Scientific has acquired NanoDrop Technologies, a leading manufacturer of micro-volume ultraviolet visible instrumentation. The acquisition strengthens Thermo's portfolio of UV-Vis spectrophotometry instruments, and came just two weeks before CEO Marijn E. Dekkers hinted at more acquisition "opportunities" to come in the wake of a strong Q3.

CLINICAL SUPPORT

CRIX International, the Clinical Research Information Exchange, has chosen Northrop Grumman to provide support and security services for the CRIX technology platform. CRIX hopes to offer life sciences organizations with services to make clinical work processes more efficient. Northrop Grumman will handle hosting, security, scaling, credentialing of users, and the help desk for the platform, which is set to be operational in early 2008.

CANCER ARRAYS

A study published in the *Proceedings of the National Academy of Sciences* (October 2007) identifying proteins associated with ovarian cancer relied on Invitrogen's ProtoArray technology to identify proteins present in cancer patients' blood that may be useful in developing future diagnostic tests. The arrays identified more than 94 candidate biomarkers with enhanced reactivity in cancer patients.

Cambridge Healthtech Institute's 6th Annual

Bio·IT World Conference & Expo

April 28 – 30, 2008 • World Trade Center • Boston, MA

Keynote Presentations by:

Linda Avey, co-Founder, 23andMe, Inc.

Joshua Boger, Ph.D., President & Chief Executive Officer, Vertex Pharmaceuticals, Inc.

John Reynders, Ph.D., Vice President & Chief Information Officer, Johnson & Johnson, Pharma R&D

Make sure to mention keycode U02 when registering!

Concurrent Tracks:

1. Informatics and IT Infrastructure & Operations
2. Informatics for Genome Analysis, Biomarkers, and Target Discovery
3. Predictive and *in silico* Science
4. eChemistry Solutions
5. Clinical and Medical Informatics

A Must-Attend Event for the Life Science IT & Informatics Community

Event Features:

- **Access** All Five Tracks for One Price
- **Network** with 1,500+ Delegates
- **Hear** 85+ Technology and Scientific Presentations
- **Choose** from Four Pre-conference Workshops and the Oracle Life Sciences and Healthcare User Group Meeting
- **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 2008 Benjamin Franklin Award and the "Best of Show" Awards
- **View** Who's Who & What's New in the Exhibit Hall
- And Much More!

Official Publication: Bio·IT World

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

OSINT and the Pharmaceutical Enterprise

JAMES GOLDEN

In its July 2004 report, the 9/11 Commission recommended the creation of an "open-source" intelligence agency — somewhat different than the CIA and NSA. Open Source Intelligence (OSINT) is defined by the Director of National Intelligence as intelligence "produced from publicly available information that is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement." OSINT focuses on creating actionable intelligence from public information, allowing other Federal agencies to focus on creating primary intelligence from covert human sources or listening in on electronic communications.

Many OSINT resources are available via the millions of Web pages, blogs and databases on the World Wide Web. Utilizing these open sources in an actionable way to produce quality intelligence requires unique research production processes and advanced search, retrieval, discovery, characterization and analysis software. I believe we have a similar opportunity within the world of biopharma — much of the information needed to create sophisticated intelligence and analysis regarding drug discovery and development, regulatory, and sales and marketing projects is available through OSINT.

Obviously, information technology is essential to the pharmaceutical enterprise. However, IT systems to help create meaningful pharmaceutical OSINT in a meaningful way have not kept pace with more process-driven information systems. Every pharma company has invested in tools and techniques for enterprise-wide search, document discovery and management, knowledge management and business intelligence, and possibly even semantic-web type tools such as ontology creation and text mining. Such tools can play an important part in creating a pharmaceutical OSINT system to better inform company decision makers.

However, technology is only one piece of the puzzle. A more central issue is knowing what questions to ask and accurately determining what constitutes an answer.

Question asking (and answering) is a fine art. In our work as consultants, we're often asked a variety of questions, from "What does Wall Street think of our CEO?" to "How many biomarkers can we actually use as prognostics in our clinical trials?" Answering each of these questions requires a different approach (we're fortunate at SAIC to have decades of National Security Intelligence experience to call upon).

In the upcoming months, I'll explore some of the concepts that are useful in building OSINT systems for drug intelligence. Here, I focus on two important topics: mining the Deep Web and taking advantage of developments in Web 2.0.

Deep Web Mining

To create relevant intelligence from OSINT sources, analysts require a well thought-out process for querying, discovery and analysis, as well as access to software, systems and processes to collect and monitor open sources for relevant information around biopharma business practices. These sources include information regarding genomic targets and biomarkers; disease epidemiology; competitive intelligence; clinical trial information including trial design, endpoints, and enrollment statistics for themselves and their competitors; business intelligence surrounding suppliers, vendors, distributors and partners; regulatory standards and recommendations; sales and marketing data (including physician script data); and investor relations and market sentiment.

BRENDAN SMIALOWSKI/AGENCE FRANCE PRESSE

To discover content on the Web, search engines typically use web crawlers that follow hyperlinks. This technique is ideal for discovering resources on the surface Web, but is often ineffective at finding Deep Web resources. (The Deep Web — or Deepnet, invisible Web or hidden Web — refers to WWW content not part of the surface Web indexed by search engines.) For example, these crawlers seldom find dynamic pages resulting from database queries due to the infinite number of queries that are possible.

Deep Web resources may be classified into one or more of the following categories:

- Dynamic content: dynamic pages returned in response to a query or accessed only through a form.
- Unlinked content: pages unlinked from other pages, which may prevent Web crawling programs from accessing the content. This content is referred to as pages without backlinks (or inlinks).
- Limited access content: sites that require registration or restrict access to their pages (e.g., using the Robots Exclusion Standard), prohibiting search engines from browsing them

and creating cached copies.
- Scripted content: pages only accessible through links produced by JavaScript and Flash.
- Non-text content: multimedia (image) files, Usenet archives and documents in non-HTML file formats such as PDF and DOC documents.

Creating a complete OSINT picture for drug portfolio intelligence requires access to Deep Web content. Several vendors supply Deep Web mining applications as part of an overall search strategy. One product I particularly like is Deep Query Manager from BrightPlanet (www.brightplanet.com). I've had good luck using their product to collect data from hard to reach places on the Internet.

Web 2.0

"Web 2.0" refers to a perceived second generation of Web-based communities and hosted services such as social networking sites, wikis, and folksonomies that facilitate collaboration and sharing between users. Though the term suggests a new version of the Web, it does not refer to updated Web technical specifications, but to changes in the ways systems developers have used the web platform.

Alluding to the version-numbers that commonly designate software upgrades, the phrase "Web 2.0" hints at an improved form of the World Wide Web; advocates suggest that technologies such as blogs, social bookmarking, wikis, podcasts, RSS feeds (and other forms of many-to-many publishing), social software, and online Web services imply a significant change in web usage.

Web 2.0 can also refer to the transition of web sites from information silos to sources of content and functionality as well as a social phenomenon embracing an approach to generating and distributing Web content, characterized by open communication, decentralization of authority, freedom to share and reuse, and "the market as a conversation."

This is an intriguing idea considering how isolated most research labs tend to be. While most drug discovery researchers tend to collaborate, the number of nodes in those networks tends to be small.

The use of Web 2.0 technologies to enable on-line communities of interest and social networking is critical to OSINT analysts. These communities allow users of varying interests to connect, network, communicate and publish content on many topics, including several that would be relevant to drug industry best practices, portfolio valuation, and related technologies. Web communities such as MySpace, Friendster, and especially scientifically focused communities such as SciLink are important OSINT sources in collecting data for biopharmaceutical intelligence.

Next month: How to tie these approaches together to create a Biopharmaceutical OSINT analyst toolkit.

Jim Golden is CTO at SAIC. He can be reached at goldenj@saic.com.

Sunny Skies for Compute Cloud

MICHAEL CARIASO

Buying CPUs by the hour is back. Marketing apparently decided that 'On Demand' sounds pushy and 'Utility' and 'Grid' are too rigid. This time it will be called a 'Cloud', and it will be running inside Amazon.com. Many groups have offered CPUs by the hour over the years, but the newest efforts seem to address many of the weaknesses of previous efforts. In August 2006, Amazon introduced the EC2 (Elastic Compute Cloud). In recent months, Microsoft began talking publicly about a "Cloud OS," and a Google-IBM alliance announced a "Cloud" academic-only collaboration.

So why now? Multicore chips may have arrived, but with the exception of a few well-established niches, multi-threaded and multi-process codes to utilize this hardware effectively are going to take a while to appear. At the same time, virtualization is maturing. Running virtual machines is a reasonable way to get a lot of use out of a multicore system. The back room and the data center have known this for years, and consumer products like Parallels and VMWare have pushed that same technology onto the workstation. More users are getting used to the idea of creating machines on demand.

Creating virtual machines somewhere across the Internet is the next step, and that is what Amazon's EC2 does. While there is still a learning curve, the adoption process is far simpler than what we've seen from others. Previous offerings have required an upfront estimate of how many hours to be used, and offers of hundreds of free hours smelled of strings and a sales rep.

Amazon charges ten cents for each hour a virtual machine is running. I can cancel or pause at anytime. Coupled with their storage network, it is possible to use precisely the amount of CPU and disk that I need with a remarkably simple setup process. Two minutes after I request my first machine, I'm SSHed into a root shell. My first hour of kicking the tires cost 11 cents. This freedom to experiment makes an enormous difference.

The machine is a virtual instance, with 1.7 Ghz x86 processor, 1.75 GB RAM, 160 GB local disk, and 250 Mb/s of network bandwidth. This costs $0.10/machine/hour. For more horsepower, machines with 4x or 8x specs are available for $0.40 and $0.80. Persistent storage is independent of the virtual machines for $0.15/GB/month. Transfers within the cloud are free, with data moving between the cloud and the internet costing $0.10/GB on the way in and $0.18/GB for each transferred out. With larger volumes, outbound transfers fall to $0.13/GB.

Clustered Pricing

This pricing model has some interesting properties. Shared resources such as a bio-mirror.net site inside Amazon's S3 storage would be fast and free to anyone inside EC2. More dramatically, consider a large parallel task with a small input and output, such as MrBayes. If the compute will take 10 machines 10 days, inside Amazon EC2 this compute costs $240. Instead the user can request 100 machines, get the results back in a single day, and it will still cost $240. Rush jobs don't cost extra. If you'd like to crunch the numbers, Amazon provides a calculator.

Amazon has also simplified is the creation of custom virtual machines. The machine I boot can be one of several standard "base" boot images (fedora, debian, Windows Server 2003, ...). I've got root so I can install software, create accounts, and make it suit my needs. After taking a snapshot of that machine, I can boot this custom system rather than the default one. This puts me into my customized environment in 2-5 minutes.

I know I should be making offsite backups of my personal files, but it was always too much of a hassle. Now I use Jungledisk.com and s3sync.net to backup my personal machines into the cloud. Services like Jungledisk are possible because Amazon handles the billing while allowing 3rd party developers to add services and surcharges. This promises to create an interesting new software service environment.

When I first saw the Watson and Venter genome sequence fasta files sitting on the NCBI ftp server I was excited. I booted a machine in the cloud, and began the downloads. During the transfers I built a small BioPerl pipeline to run mpiBLAST and CLUSTALW. While making a Waston+Venter blastable database, I downloaded dbSNP. When that completed I booted a few more machines, and crunched the data for a few days. Afterwards I copied the results down to my local machine, and then turned off my little cluster. I'd spent less than $50.

Facilities that produce large volumes of raw data will continue to find the need for onsite compute resources. Network pipes to the laboratory will not keep pace with next generation sequencing technologies. Some facilities may incorporate a remote cloud into their schedulers. For the moment, many will find that no matter how compelling the service might be, they cannot bring themselves to ship their data offsite. These folks will continue to build out ever larger machine rooms. But if you don't produce primary data, you don't need a data center. Smaller shops working with data from the public domain or remote collaborators, including the next generation of biotechs, may find little reason to maintain onsite servers.

I suspect a few network admins will view these as storm clouds, but for science they appear to have a silver lining.

Michael Cariaso is the senior scientific consultant for the BioTeam. He can be reached at cariaso@bioteam.net.

Back to Omics

MALORYE A. ALLISON

"When we start to think about the biology of disease, our goal is not just new targets anymore but biomarkers also," said Andrew Plump, executive director of Merck's Cardiovascular Disease franchise. Plump was speaking last month at the Personalized Medicine meeting held at George Washington University's Richard B. and Lynne V. Cheney Cardiovascular Institute, in Washington D.C.

The meeting was focused on "The Genomic Revolution in Cardiac Care," and highlighted how biomarker research is now not just reaching across a wider range of conditions, but also using a broader mix of tools. Even more intriguing, omic tools are increasingly being used together to uncover valuable biomarkers to guide those now "must have" proof of concept studies bridging pre-clinical and clinical trials.

Not long ago that was deemed too expensive, too complicated, and unnecessary.

Finding biomarkers is particularly challenging in fields such as atherosclerosis, where biopsies are not a routine part of patient monitoring. It's hard to figure out what's happening to all that nasty arterial plaque the medications are supposed to whisk away.

Merck has a leg up because it has been a pioneer in omic analysis, dating back to a high profile acquisition of subsidiary Rosetta Inpharmatics in 2001. "We bought Rosetta in part for their real capacity in system analysis," Plump said.

In trying to uncover blood-based biomarkers of atherosclerosis progression, Merck researchers are examining SNPs, histology, proteins, and RNA. They are not just looking for genetic causes of cardiovascular disease, but for markers of "inflammatory instability" as Plump said. "Right now, there are no tools for decision making in phase 2," said Plump. "In the future, imaging tools will emerge, but in the meantime we are looking for markers in plaque."

Not all pharmaceutical companies have the internal omics capability to find such markers by themselves, however, and that's created opportunity for a new wave of technology vendors and service providers with specialized expertise or novel tools.

Digilab Biovision falls in the latter category. Offering a novel "peptidomics" service, the company's tools uncover peptide and small-protein based biomarkers. Because they are signaling molecules, peptides are believed to be a more specific read out than proteins or RNA, says Hans-Dieter Zucht, chief technology officer. And what have we learned about analyzing omics so far? "The sequence does not speak to you," warns Zucht. Experiments must be appropriately designed, and researchers need to get enough samples. In some cases, Zucht says, his company and their collaborators are aiming for as many as 10,000 samples to power a single study.

After sifting through all that data, everyone hopes to find just one golden marker in the end, but that's not always possible. "Especially when you are looking at markers for a syndrome, such as inflammation," Zucht says, "you're probably going to need more than one marker."

Tangible Progress

Proteome Sciences' business development director Ian Pike adds another bit of hard-earned wisdom. "You have to find the markers and identify the proteins," he says. Researchers working in proteomics have slowly shifted away from using random sets, or constellations, of unknown markers, to wanting to know exactly what proteins make up the marker. Proteome Sciences has developed novel tags and other methods that make it easier to quantify proteins and measure them in complex samples.

Few of these tool providers can parade big pharma customers who have used their tools, the field is too secretive. But

> The sequence does not speak to you. Experiments must be appropriately designed, and researchers need to get enough samples.
>
> **Hans-Dieter Zucht**, Digilab Biovision

Jack Reynolds, formerly a senior vice president in R&D at Pfizer, speaks generally about Genstruct, which has a platform that is both novel and specialized. Genstruct generates and analyzes data from metabonomics, proteomics, protein phosphorylation, and gene expression.

"People didn't get that much out of genomics originally," Reynolds says. "Because you have to create a baseline platform and generate knowledge from it, like Genstruct has done." Reynolds says this platform, which Genstruct CEO Keith Elliston says has been used in about 30 partnerships so far, helped Pfizer become one of the first companies to submit to the FDA under the voluntary genomics data submission guidance.

"In one area of adverse events we submitted over a million data points," Reynolds says. "People would ask how can you get something out of that much data, but that's exactly why you need something like the Genstruct platform."

Clearly, tangible progress has been made, but pioneers like Merck, Digilab Biovision, Genstruct, and Proteome Sciences are all still working to gain further understanding of the how these data types interplay and what types of conclusions can, and cannot, be drawn from them.

Sunlight and Systems Biology in Seattle

ISB course whets appetites and spurs projects.

BY JOHN RUSSELL

Seattle — On this Monday in July it is not raining in Seattle. Apparently this is typical in summer and contrary to the region's dreary image. A record-setting heat wave had bikers and joggers out in force under bright sun. Inside the Institute for Systems Biology (ISB), thirty or so "students" gathered for the first day of ISB's intensive week-long introduction to systems biology course and I was one of them.

It was a diverse group, drawn heavily from academia. About half had Ph.D.s (or soon will) in something and their experience ranges from newbies and post-doc (several) to experienced researchers (somewhat fewer). The rest had varying backgrounds. Gautam Venkatesan, a Stanford University-trained economist with experience at Hewlett-Packard and a software startup, ended up asking many of the week's most interesting questions. There was a high school biology teacher from Tacoma, a medical lab worker from Dubai, and, of course, a science journalist (someone has to drag down the curve).

It was also a decidedly international group. There were attendees from institutions in Sweden, Italy, Switzerland, Spain, Qatar, and the United States. For five days, we were immersed in the concepts and tools that comprise systems biology (SB). We tackled a lab experiment designed to put SB ideas and tools into practice — measuring, analyzing, developing a hypothesis, and then testing the hypothesis. By the end of the course, "Each student should be able to develop a research outline that could form the core of a systems-based grant application," say course materials.

That seemed like a stretch for me, but not for more advanced participants. In fact, one attendee, Scott Berceli, has already done just that. Berceli is a vascular surgeon affiliated with the University of Florida who also has Ph.D. in chemical engineering.

ISB's John Aitchison says systems biology is powerful, but can't perform miracles.

"I attended the course with a specific purpose," he said, "to realign some of my multidisciplinary research activities with some of latest concepts in systems biology. In particular, the discussion of gene regulatory networks, and mathematical approaches for solution of these complex systems, provided me an important step forward in re-tooling my ongoing research efforts."

Systems Projects

No stranger to molecular biology, the course proved broadly useful to Berceli but perhaps more valuable were independent meetings with ISB founder Leroy Hood and other investigators at the ISB, which provided a cornerstone for an NIH/NHLBI systems biology proposal Berceli submitted in September 2007.

Berceli's ambitious proposal, "examines vascular remodeling through a multi-scale approach, conceptualizing the process of vascular adaptation as two parallel, but interconnected, processes. The global remodeling response is mediated through variations in the gene regulatory network, while the focality of lesion development is modulated through the dynamics of monocyte homing to regions of altered flow."

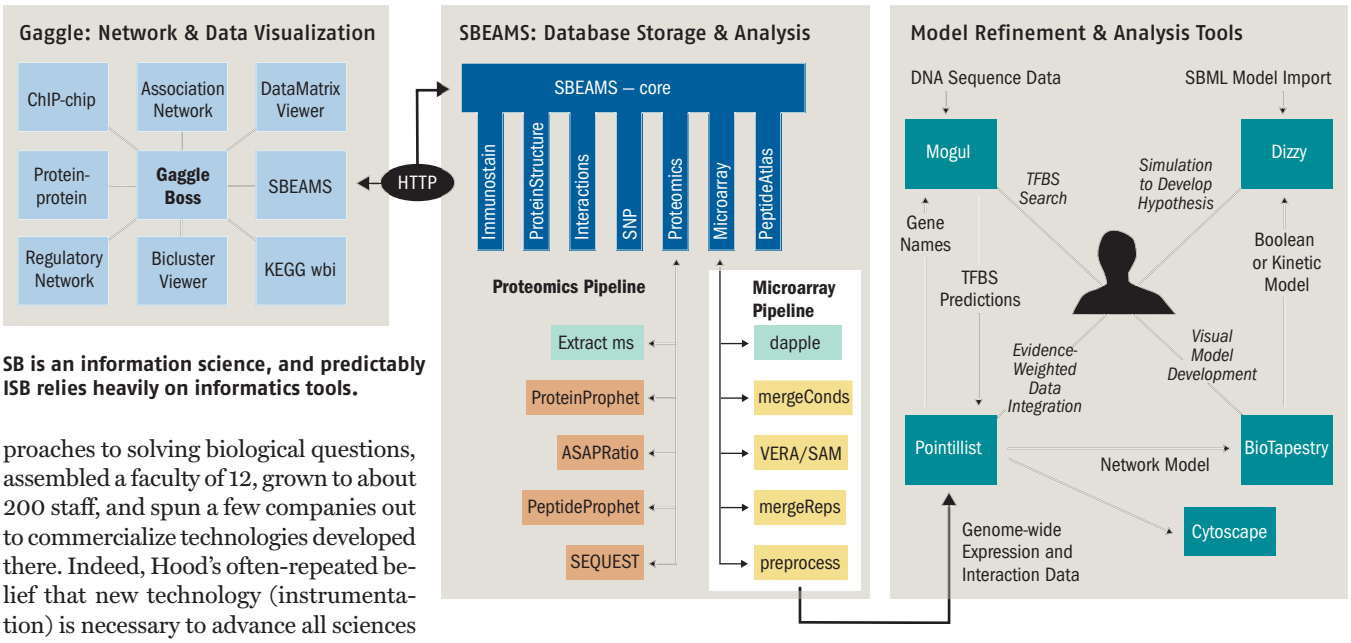
My goal was more modest: to have a hands-on experience (lecture, lab, and informatics tools) that crystallized the ideas, goals, and future directions of systems biology. Berceli was part of my lab group, and his guidance helped ensure my unpracticed lab skill didn't muck up our piece of the class experiment. Indeed one of the cornerstones of systems biology is multidisciplinary collaboration and the 3- and four-person lab groups we divided into delivered a bit of that experience.

No doubt ISB has many goals in offering the course. Planting SB project seeds was certainly one, and that's happening to some degree as exemplified by Berceli's grant proposal. A few attendees said they had "job interviews" lined up inside ISB during the week so clearly recruiting is also going on. Identifying possible collaborators is another objective. And, of course, just getting the systems biology word out is important. The cost is fairly modest, $2000 for corporate attendees and $750 for students and postdocs. (ISB waived tuition for me.)

The opportunity to directly interact with Hood and prominent ISB staff was clearly a big attraction. Hood kicked off the week with an hour-long discussion of systems medicine and offered his definition of systems biology (See, "ISB is a Sure Cure," *Bio•IT World*, Aug. 2007). Hood's vision is powerful and his track record impressive. He takes aim at biology's past dependence on rote memorization and cites his early training as an undergraduate at CalTech. He had Richard Feynman for freshman physics and Linus Pauling for freshman chemistry. Nice lineage! Both emphasized conceptual clarity and quantitative approaches, says Hood, and that's stuck with him — the essence of systems biology.

Hood along with Reudi Aebersold and Alan Aderem founded ISB in 2000 to help chart systems biology future. Since that time ISB has broken a good deal of new ground in systems biology (particularly in technology development), championed the cause of multi-disciplinary teams and ap-

Informatics at ISB



SB is an information science, and predictably ISB relies heavily on informatics tools.

proaches to solving biological questions, assembled a faculty of 12, grown to about 200 staff, and spun a few companies out to commercialize technologies developed there. Indeed, Hood's often-repeated belief that new technology (instrumentation) is necessary to advance all sciences is a dominant theme at ISB.

Buoyed by Hood's opening salvo, the real work of the course began. Most days had two components, class and lab (wet and informatics). We started promptly at 8:30 AM and finished around 4:30 PM. ISB staff provided lectures on all things SB, including proteomics, functional genomics, network analysis, modeling techniques, systems genetics, hypothesis generation and testing, etc. Much of the material is readily accessible, but instructors often dig deep and are generally willing to have follow-up discussions later.

SB in the Lab

The lab exercise is designed to demonstrate an SB approach. We attempted to learn how a salt-loving bacterium, Halobacterium NRC-1, responds to copper stress. We did some cell culturing, took samples at different times, extracted RNA, determined its quality, labeled it, ran the microarrays, processed the raw data, used informatics tools to store, visualize and analyze results, imported appropriate pathway data from public sources, developed a hypothesis, and tested the hypothesis. In other words, quite a bit.

Some of us (me) hadn't pipetted since the days of mouth pipetting! One of us (not me) cleaned a slide before the sample was applied, wiping off the probes! Our slides, prepared by ISB, had the full set of the bacterium's 2400 genes. Fortunately, we were repeating portions of an experiment ISB has previously published. And the ISB staff was well prepared for most contingencies. Perhaps most surprising was that most of our data turned out to be quite useable.

A fair amount of time and effort throughout the course was devoted to introducing basic informatics tools and, again, the range of familiarity with those tools varied widely among attendees. Each student was provided with a laptop in class, and ISB personnel were generally able to keep students on track though it was occasionally challenging.

ISB primarily uses open source tools, many of which have been developed by ISB. The institute uses a home-grown software framework called Gaggle Boss, into which individual software tools called — you guessed it — Geese are plugged. A researcher can move data fairly easily from one tool to another using Gaggle. A few of the tools we used included Cytoscape; R statistical environment; Data Matrix Viewer; and we imported data from the KEGG database.

The class tentatively identified two responses to stress including gene networks for a chelating agent and a membrane transport mechanism; we then confirmed the findings.

Course director John Aitchison emphasizes, "[The course] provides an overview of systems biology, including sufficient background to understand the strengths and limitations. Systems biology can do a lot of things, but it does not work miracles. Students [will] understand its possibilities, get a sense of the promise, get excited about it and, where possible, incorporate systems thinking and systems biology into their projects."

Venkatesan was sufficiently inspired to subsequently enroll in a genomics and proteomics course at the University of Washington, Department of Genome Sciences. "I hope to work in industry," he says. "The people at ISB are truly distinguished pioneers. It is difficult to overestimate the complexity of the data and methods that we are undertaking." It's not hard to imagine this '02 Stanford grad running a SB company and, of course, Berceli has high hopes for his grant.

On the final day, Hood again addressed the class. His broad opening and closing presentations served as bookends for our time at ISB. This time he also took considerable time to answer the class's many questions — the crowning event on what was a very worthwhile week. ●

Surviving the New Patent Regime

New USPTO rules would limit rights if passed.

BY JEN ZARUTSKIE

Last August, the United States Patent and Trademark Office (USPTO) published controversial rule changes to come into force November 1. On Oct. 31, GlaxoSmithKline (GSK) was awarded a preliminary injunction by the U.S. District Court for the Eastern District of Virgina to delay the new rules from taking effect. (At press time, the hearing date had not been set.)

As set by the USPTO, the new rules would place an additional burden on patent holders and those seeking new patents by limiting the number of continuation patent applications that may be filed and the number of claims that will be examined per application as a matter of right. GSK argued that the rule changes were "not within the legal purview" of the USPTO, and would impede innovation in the pharmaceutical industry.

Rules state that patent seekers may no longer file an unlimited number of multiple continuations to describe and claim ultimate products developed from a basic pioneering patent and may no longer expect to easily obtain multiple patents to different claim sets covering different aspects of their products, a common practice in the long drug discovery process.

If the USPTO rules are upheld in the courts, there will be new issues and strategies to consider when preparing and managing an IP portfolio in order to maximize the now more limited patent rights available. Under the rules in question, applicants may only file two continuation applications and one Request for Continued Examination (RCE). Additional patent applications must be accompanied by a petition showing that the new amendment or evidence could not have been submitted during the prosecution of the earlier application. When filing a continuation-in-part (CIP) patent application, applicants must now identify all claims supported by the parent application's disclosure. Divisional applications may only be filed if the claims were the subject of a USPTO restriction requirement that is not traversed by the applicant or based on a Suggested Restriction Requirement (SRR) filed with the USPTO. (In each divisional application, two continuation applications/CIPs and one RCE can be filed.)

In addition, no more than 25 total claims will be examined in any single application, unless a complicated examination support document is filed. The USPTO rules presume that multiple applications having the same priority or filing date count as a single application for claim number purposes if the applications include substantially overlapping disclosures, a common inventor, and a common owner.

Strategies for Managing Portfolios

For applications that have not been examined or are not subject to a restriction requirement, applicants should make sure that all disclosed inventive subject matter is claimed, adding additional claims if necessary. Then consider whether to file an SRR, await a restriction requirement, or call the Examiner to discuss appropriate restriction to the claims. The more restriction groups, the more divisionals that may be filed. Once the SRR is accepted or the restriction requirement arrives, an applicant should cancel non-elected claims and request a refund of the excess claims fees.

> There will be new strategies to consider when managing an IP portfolio.

Applicants should organize any potentially material prior art and if they are planning to perform a prior art search, do so earlier in prosecution and submit it (ideally) before a first office action. Further, applicants should realistically consider their chances of winning an argument versus amending the claims, and consider more aggressively amending in non-final office actions in order to preserve the option to file an RCE without a petition and showing. Applicants should also consider more frequently holding interviews with examiners, so that arguments and amendments can be discussed before potentially wasting an office action response. Finally, applicants should consider making more use of the appeals process in cases where the examiner is clearly misinterpreting and misapplying the law, rather than engaging in rounds of fruitless argument trying to readjust the examiner's view.

The new rules would also require rethinking how and when to file new applications, and what to include. Applicants should assume a CIP is not available. They should consider if it is "too early" to file on a discovery, and if it is possible to develop the invention further without triggering any bars to patentability, such as publication. Rather than file utility applications, applicants should consider relying on provisional applications to obtain filing dates and an extra year to develop the invention. For example, applicants may consider filing multiple provisionals around an invention as it evolves, wherein the provisionals are cumulative of the original filing.

Applicants should consider pursuing first those claim sets that are best enabled and appear to be the easiest to argue before the USPTO, the "low-hanging fruit," in a serial prosecution strategy. This will buy time to conduct research for the claims that will need to have data "sworn in" and preserve the continuations needed to pursue such claims when they will be more easily prosecuted. Applicants may be able to make this easier by combining all possible inventions surrounding a discovery into a single application, and writing the initial claim set to support the most complex restriction requirement possible.

In response to the GSK objection, the USPTO issued a statement supporting the rules, to the dismay of the biotech, pharma, and university communities. The practical consequences for drug discovery will be decided in the hearing. ●

Jen Zarutskie is an associate at Foley Hoag. Email her at: jzarutskie@foleyhoag.com.

ABCD: The Relentless Pursuit of Perfection

Johnson & Johnson's IT team built it — and they came.

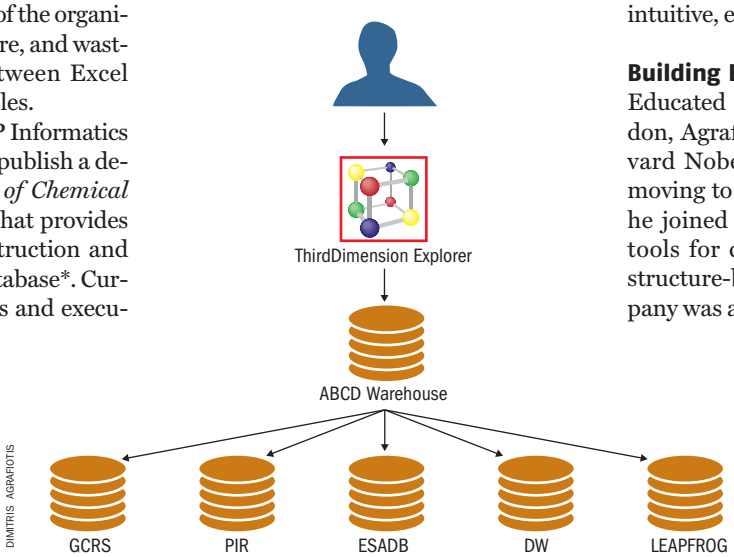
BY KEVIN DAVIES

Four years ago, Dimitris Agrafiotis' team from recently acquired 3-Dimensional Pharmaceuticals (3DP) was entrusted with building a major global informatics platform, dubbed ABCD (Advanced Biological and Chemical Discovery) for Johnson & Johnson Pharmaceutical Research & Development (JJPRD). More than many big pharmas, J&J scientists were struggling because of the organization's decentralized structure, and wasting time shuffling data between Excel spreadsheets and HTML tables.

This month, Agrafiotis, VP Informatics at JJPRD, and 21 colleagues publish a detailed paper in the *Journal of Chemical Information and Modeling* that provides the framework for the construction and applications of the ABCD database*. Currently, nearly 1,200 scientists and executives across J&J's three major research and early development (RED) sites are using ABCD. (Agrafiotis quickly checks the precise number — 1,182.) The primary data in ABCD hails from the three RED units and their satellites — La Jolla, Calif., Spring House, Penn., and Beerse, Belgium.

Agrafiotis says ABCD has changed significantly in the past few years (See, "How to Spell Discovery," *Bio•IT World*, June 2004), even since the scientific paper was submitted. "We've added *in vivo* data, we've cleaned up and harmonized a lot of the result types, and added more data feeds such as compound availability. Most importantly, we've restructured the fact tables to support more effective multi-dimensional aggregation of data at the compound and biological result levels, in order to make reports a lot more effective at all levels of resolution," says Agrafiotis.

He sounds insulted when asked if there is any commercial software integrated into ABCD. "Zero," he says. "It's all homegrown. The only commercial systems we use are Oracle on the back end and Informatica for the ETL (extraction, transformation, and loading). And a couple of component libraries for the application."



Schematic view of the ABCD discovery informatics landscape. Legacy source systems include the Global Compound Registration System (GCRS); the Pharmacalogical Information Retrieval system (PIR); the Extended Structure Activity DataBase (ESADB); the DiscoverWorks integrated informatics platform (developed at 3DP); and Leapfrog, a target and project tracking system.

Management, he says, has been surprisingly open and "extremely supportive" of the ABCD initiative and particularly of doing it in house. "It's unheard of in most other companies," says Agrafiotis, "where they either stitch together commercial packages or contract out most, if not all, of the development."

Particularly impressed with ABCD is John Reynders, the recently appointed CIO for Johnson & Johnson's R&D program. "From the outside in, I was very impressed with the platform," says Reynders. "Now that I'm on the inside, I'm even more impressed with the platform. It's something, frankly, I view as us being able to leverage more broadly across the pipeline."

The scientists seem content too. "By bringing together and integrating data sources from around the world, ABCD has fundamentally changed the way scientists from many disciplines work together," says Peter Connolly, research fellow in medicinal chemistry at JJPRD. "Think about how the Internet has changed the world during the past decade. That's what ABCD is doing for collaborative science at J&J — bringing information quickly to the desktop using an intuitive, easy-to-use interface."

Building Blocks

Educated in his native Greece and London, Agrafiotis did a post-doc with Harvard Nobel laureate E. J. Corey before moving to the pharma industry. In 1994, he joined 3DP, building computational tools for combinatorial chemistry and structure-based drug design. That company was acquired by JJPRD in 2003.

ABCD consists of two major components — a data warehouse on the backend, and the "Swiss army-knife" Third Dimension Explorer application (3DX) on the front end. There's also a workspace web portal for delivering and managing ABCD information. "The warehouse did not exist before we started this initiative," says Agrafiotis. "3DX was based on a lot of code developed at 3DP over the past ten years. ABCD as a goal was in management's minds when 3DP was acquired, but the implementation had not been decided."

Agrafiotis points out that the switch to ABCD has been completely voluntary. Management did not twist anybody's arm to use the new system. Most of the legacy systems are still in place and operating, but their usage is very light.

*Further Reading: Agrafiotis DA *et al.* "Advanced Biological & Chemical Discovery (ABCD): Centralizing discovery knowledge in an inherently decentralized world." *J. Chem. Inf. Model.* November 2007.

Computational Development

Most of the users fall into three groups: computational scientists, project champions and data managers, and technology-savvy scientists. Today, scientists conducting chemical queries routinely go through ABCD, using the query wizard embedded in 3DX. This allows them to directly extract data from the database. Some might also use tools from Scitegic (such as Pipeline Pilot) to manipulate the data. "The query interface allows them to combine chemical and biological queries, search by site, by chemist, by project — they can do it all through 3DX," says Agrafiotis.

He adds that in Beerse, Belgium, JJPRD scientists are using Project Datasets — pre-packaged queries assembled from ABCD and updated on a daily basis. "These ideas are now permeating across the organization," he says. The database is primarily used in drug discovery research, but its use is expanding into other functional areas. For example, Centocor, one of J&J's biotech subsidiaries, is using ABCD for clinical data mining purposes.

Agrafiotis manages an RED IT staff of about 36 people, with a dozen or more people working on various projects under the ABCD umbrella. The team manages a support forum for all users and developers. If anyone has a problem or issue, they can post a message to the forum. ABCD staff receives tens of messages each day.

There is also a scientific advisory team composed of "influential internal scientists who have a passion for informatics," says Agrafiotis. The IT savvy representatives represent all the sites, and meet monthly — more often if there is an urgent issue. Typical issues the advisory team deals with might include: How to score assay results so that users can gauge whether a compound is good, bad, or indifferent? How to normalize data? How to classify *in vivo* protocols?

The front end of the ABCD system, Third Dimension Explorer, 3DX, helps visualize data.

Competitive Advantage

Although hardly impartial, Agrafiotis is convinced that the resource affords J&J a competitive advantage. "Absolutely, no doubt about it," he exclaims. "Otherwise management wouldn't support this effort so strongly. Support hasn't wavered at all, in fact it has intensified." Connolly agrees: "To my knowledge, there's nothing in the pharma industry to equal its scope and usability," he says.

"Although it's had a cheminformatics focus in the past, what I've been very delighted to discover is the architecture is very general," says Reynders. "It's built on SOA and is a multi-tiered environment — it is very easy to add plug-ins and capabilities."

Other companies are grappling with the same kinds of issues, says Agrafiotis. "But frankly, it takes a unique ability to do this themselves, the ability to assemble a disciplined and effective team for this type of initiative to succeed." Everyone in his team "understands we need coherence and consistency in the way the interfaces are built, in the way the system performs. It's like the old Lexus tagline: 'The relentless pursuit of perfection.'"

Agrafiotis cites an example. He was working with a modeler in La Jolla on developing a chemical cartridge prototype. "I wanted to search the entire database in less than a second," he recalls. "I had the prototype down to a couple of seconds, and it beat the heck out of everything out there. But we didn't settle for that, we wanted to be able to do it in sub-second timeframes."

A few days later, it was done, and Agrafiotis says the rest of his team is wired the same way. "The team on ABCD has an obsessive drive for excellence in every single aspect of the project: the application, the database, the overall user experience. We look at every piece of code as a statement about our capabilities."

The recent hiring of Reynders "will be very positive," predicts Agrafiotis. Adds Reynders: "I'm looking forward to working with Dimitris to extend it into other areas of informatics — bioinformatics, translational informatics, medical informatics, so it can become a platform that extends across the pipeline."

Not one to rest on his laurels, Agrafiotis is busy making plans for the next phase of the database. "The first release was about downloading data, visualizing, mining, analyzing it," he says. "ABCD is now well into its second phase — the reengineering of all transactional systems — which will allow the biologists to register data into the database."

The existing systems, he says, leave much to be desired, as they use old technology, differ from site to site, and use different interfaces: "We're doing this with the same kind of philosophy and with 3DX as the primary user interface. The same application we use for data analysis and visualization, we also use to fit IC50 curves and upload the results into the database, to manage plates, to register and track reactions performed by the chemists so you don't need a paper notebook."

In short, "a coherent, unifying solution." ●

[GUEST COMMENTARY]

'Cross-Omics' and Systems Toxicology

Systems approaches combine toxicogenomic data with toxicological endpoints.

BY KURT ZINGLER

Some of the latest and most intriguing advances in biomarker development are emerging in systems toxicology — the combination of traditional toxicology methods with new strategies and tools for integrating high-throughput transcriptomics, proteomics, and metabolomics data. The goal is to better understand and predict potential toxicities at an early stage of drug development, so that biopharmas can gain deeper insights into the biology underlying toxicity, and make "go/no-go" decisions well before committing to further development and clinical trials.

The great challenges of developing effective systems toxicology approaches are well understood: making productive use of the disparate resources available as well as the acceptance of new technologies. While toxicologists have a mass of data that could help them better understand and predict toxicity, they are often saddled with incompatible data types, formats, databases, and analysis tools. They are surrounded by information that might bring them forward, yet they lack the means to make use of it, and derive informed decisions in compound profiling programs.

The emerging solution is the creation of a common system that can capture classic toxicological analyses as well as more recent "omics" data, and then provide a framework to easily make use of these data to analyze, relate, compare, and share this information — from bench researchers all the way up to top-level decision-makers.

In addressing the core technological issue of this endeavor — namely integration of toxicogenomic data with conventional toxicological endpoints — researchers face several technological and methodological limitations. Specifically, they frequently lack:

- Efficient workflows that capture, store, and analyze toxicological data, based on integrated systems;
- An ability to make decisions based on integrated information that captures the breadth of the available data;
- Access to solutions or reports that do not require end-users to be informatics experts.

The key question at this stage of systems toxicology research is, "What will it take to create a truly integrated approach?" Evidence derived from a growing number of collaborations between biopharma groups and bio-IT vendors shows that researchers require support in four main areas: integration with major high-throughput technology platforms including objective data quality assessment; integration capabilities across varied data types and formats; new standardization and automation methods for processing and analysis; and full-spectrum workflow, from sample stage to final study and result reporting.

> The goal is to better predict and understand potential toxicities at an early stage.

Processing Tox Data

In a typical scenario, toxicologists use a wide range of data from conventional tox analyses (enzyme assays, histopathology, animal observations, etc) and more recently, genomic data (transcriptomic, proteomic, and metabolomic).

While each of these data types may help characterize and predict certain types of toxicity, the data stores and analysis tools for each information source are fragmented. The problem is compounded by disseminating the information through specific program groups and decision-making teams. While higher-level committees can view static information, they lack the analysis, visualization, and interpretation tools to do additional comparisons and derive educated conclusions.

The required solution is a unified system where toxicologists and program committees can review and assess data across a wide range of technologies and models to facilitate knowledge-based compound promotion. An ideal approach provides a common framework and analysis system for the toxicologist and subsequent review committees. Each can look at the breadth of available data, drill down to specific details and easily import additional data.

As a sign of the importance of new technologies for toxicological assessment, a number of biopharma/bio-IT consortia are looking to improve the process of compound progression. Examples include the InnoMed PredTox Consortium in Europe, the U.S.-based Critical Path Initiative's Predictive Safety Testing Consortium, and the Japanese Toxicogenomics Project. Collectively these consortia are examining a variety of systems toxicology approaches to address the field's key challenges.

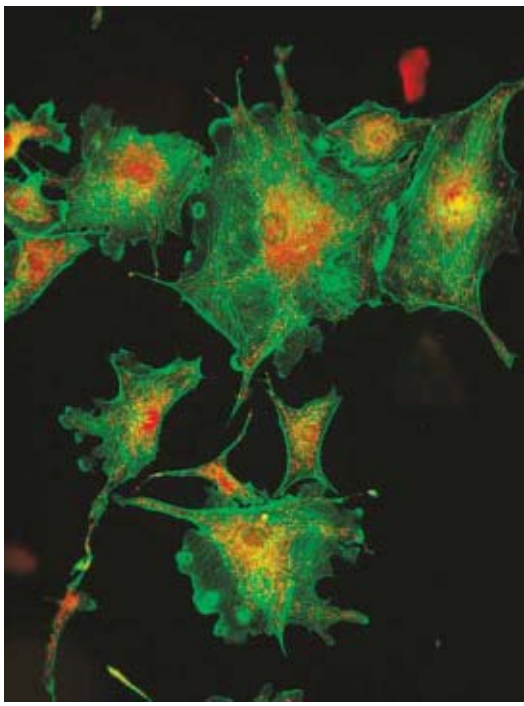
For example, InnoMed PredTox is looking at a range of compounds that failed late in the process, to determine if newer technologies (or combination of technologies) might have allowed an earlier, less costly decision to halt or redirect development.

The consortium has helped develop sophisticated databases (such as the InnoMed PredTox database) and data analysis systems, while at the same time providing validation for these emerging technologies and analysis strategies and education of scientific experts and regulators.

The ultimate goal of the consortia is to provide researchers with the breadth of available data in a shared, easy-to-use platform so that organizations can better understand the biology underlying toxicity and make scientifically justified decisions about compound progression. ●

Kurt Zingler is head of U.S. Business at Genedata. He can be contacted at kurt.zingler@genedata.com.

HIGH CONTENT ANALYSIS



As HCA matures, developers walk a thin line between power and ease of use. By Jim Kling

High content analysis (HCA) — also known as high content screening — makes possible massively parallel experiments that can reveal much about the inner workings of cells and their response to stimuli such as drugs or signaling agents. Typically, an agent or cellular proteins are tagged with a fluorescent marker that can then be imaged.

Many processes can be studied using HCA, including intracellular translocation of proteins; movement of proteins in response to activation of a receptor or a cellular pathway; and protein co-localization. Such studies have enormous potential to streamline drug discovery. For example, HCA has been used to identify or validate targets using RNA inhibition, and in secondary screens to detect cellular signs of toxicity. It can also provide visual evidence of a cancer agent as it blocks cell division, thus providing mechanistic clues. Studies that validate a drug target or identify toxicity in the earliest stages of preclinical development could drastically reduce the nearly $1 billion that is typically sunk into the development of a novel drug.

Nevertheless, early predictions that HCA would revolutionize drug development have not entirely panned out. Many pharmaceutical companies have set up HCA or HCS programs, but they have not been fully embraced — partly because it hasn't yet been proven that high content approaches can streamline drug discovery enough to justify the investment. Another important issue is that software systems that control the instruments can be difficult to learn.

WALKING A THIN LINE

HCA instrumentation has matured to the point that systems of comparable acquisition speed, magnification, and resolution are difficult to distinguish from one another because the optics have been so well optimized. Hence other factors typically drive purchasing decisions, and software is perhaps the most important. "Now that we're almost feature complete on imaging systems, our focus needs to be totally on software for the foreseeable future," says Jan Hughes, general manager and vice president of bioresearch for Molecular Devices.

HCA software developers must walk a thin line. On the one hand, users want powerful software that can perform just about any analysis they can think of. On the other hand, they want it to be easy to use. Industry experts agree that ease of use is critical. "If it's only the province of the highly talented core laboratories and can't be done by bench scientists, it has the

potential to retard" the adoption of HCA, says Mark Collins, senior marketing manager for cellular imaging at Thermo Fisher Scientific, which acquired industry vanguard Cellomics in 2005.

It isn't always easy to convince customers to scrutinize software when making a purchasing decision. "We find that probably the most difficult challenge is going to a novice imaging customer who wants to get into high content analysis or high content screening, and explaining to them the importance not just of image acquisition, but the database, image analysis, and informatics tools. They're focused on the hardware. As an end user, software can constrain you or be enabling. You better worry about your software just as much as your hardware," Hughes says.

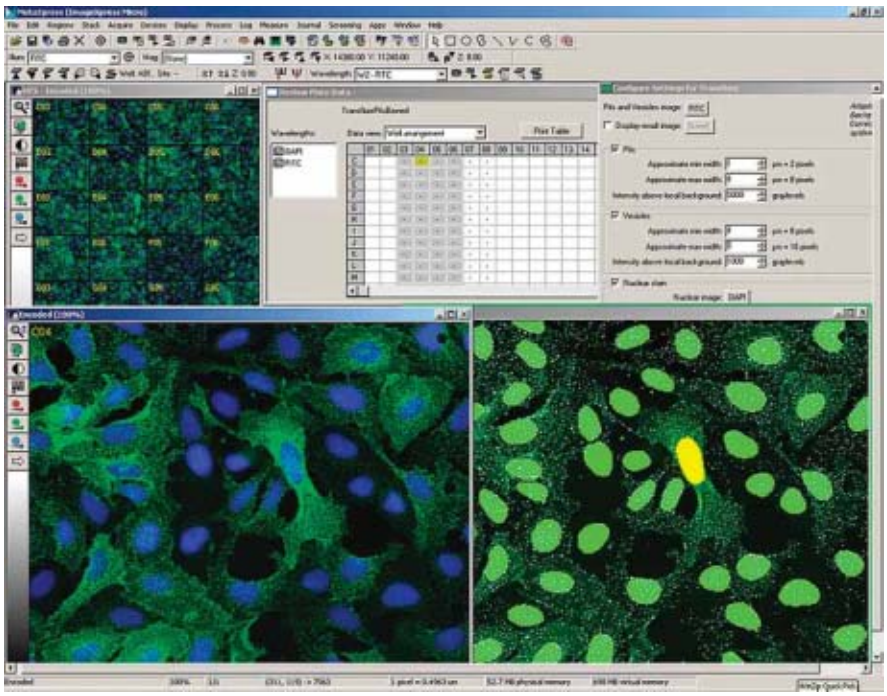
Vendors are taking various approaches to simplifying software while maintaining complexity, though it is a tricky tightrope to walk. Companies have taken various steps toward this goal. Thermo Fisher has streamlined the naming conventions for its parameters and measurements. Users click on a cell of interest and use its characteristics to create a training set for modifying the assay parameters — for example, to set a threshold for nuclear area.

The company is also continuing to develop 'out of the box' protocols for key biological processes that can be combined with Thermo Fisher's HCS Reagent Kits. Evotec Technologies, a former subsidiary of Evotec AG now part of PerkinElmer, has reduced the number of modules in its Acapella image analysis software by automating some of the functions, such as cell nucleus detection.

Other related functions have been bundled into a single module. "That makes it very flexible on the one hand, but also very user friendly," says Martin Daffertshofer, leader software development, Molecular Medicine at PerkinElmer. Evotec Technologies also offers about a dozen "canned" solutions for studying GPCRs, kinases, and other specific systems.

GE Healthcare has tried to boost usability by designing its Investigator software to operate like an interactive web page, with drop boxes, descriptive text, and very few parameters required to run an assay. "You can apply these software modules without any training," says Jacob Tesdorpf, marketing manager platform software for GE Healthcare.

Molecular Devices has two versions of its image analysis software, one geared toward research applications, the other toward high throughput screening. Both versions of the software use the same underlying engine, but MetaMorph is geared toward research applications, while MetaXpress is aimed at imaging high throughput screening experiments. Both software packages include many of the same functions, but they have interfaces that are tailored for the end user. MetaXpress development will continue to specifically address the need for fast analysis and the management of large amounts of image-derived data, Hughes says.



Molecular Devices' MetaXpress software aids researchers in visualizing high throughput screening experiments.

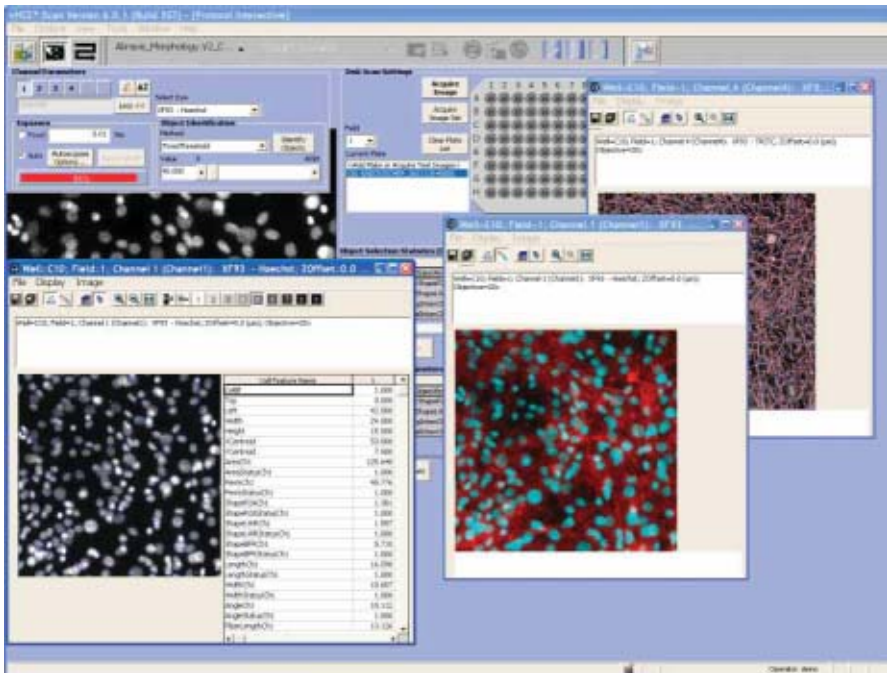
Definiens has taken a similar approach. Its Enterprise Image Intelligence Suite has several components that can be purchased individually, each to be used on a different rung of the research ladder.

For example, there is a viewer client that allows a manager to view results without influencing the experiment. An analyst client allows a technician to run preprogrammed methods and change parameters. The architect contains pre-built modules that allow a user to construct a novel assay. "They map to different work flows," says Kurt Scudder, field applications scientist for Definiens.

Thermo Fisher's Collins admits that the challenge of balancing flexibility and power with ease of use continues to be a major challenge. "I don't think we've got it yet," he admits. "I don't think anyone has."

Ease of use versus flexibility isn't just limited to choice of software. Another option is to split tasks between different instruments. For example, Genetix has developed Cell Reporter, a high content screener with an emphasis on speed and ease of use. Such machines can serve as a first run analysis. "You can whittle [a couple of hundred lead compounds] down to 10 or 20, and then take those into the machines with higher magnification [like the GE InCell or Evotec Opera]. You can reduce your bottleneck earlier in the drug discovery process," says Nicol Watson, product manager for CellReporter software at Genetix. "You don't need more than 15 or 20 minutes to train on it." (See "Tools for Therapeutics.")

Ease of use is one thing. Compatibility is another. It's not uncommon for high con-



Thermo Fischer Scientific vHCS Scan imaging software features streamlined naming conventions so that users can view a cell of interest with all of its characteristics.

tent analysis and high content screening labs to use more than one type of instrument, and that creates problems. "Screening labs will often have three (or more) different imagers, each with different characteristics. It's not uncommon at all for a screening group to have multiple platforms that can't talk to one another. That [communications blackout] is changing, but at a glacial pace," says Definiens' Scudder.

Definiens' software is designed to run on any system, giving it an edge over built-in software packages. "Where we come into play is if you want to develop a new assay every month," says Scudder. "You need an environment where you can develop your own capabilities. Our system allows you to develop completely new assays and attach them to any instrument on the market. If you want to run the same assay on a Cellomics Array Scan as on a GE InCell 1000, you can use our software to do the same analysis." Adds Mark Watson, head of life science marketing at Definiens: "You don't have to worry about comparing apples to apples."

There are efforts underway to standardize the data formats for HCA experiments, which should make it easier to swap data between instruments. Perhaps the most important element of such standards is the so-called metadata. It includes experimental parameters, such as the magnification, the nature of the dye used in the experiment, exposure time, and other factors. "That information is critical to analyzing those images and also interpreting the data from that image analysis," says Collins.

To that end Thermo Fischer has taken it upon itself to develop a standard it calls MIAHA, for 'Minimum Information for a High Content Assay.' It's a combination of the OME (Open Microscopy Environment) standard commonly used in the microscopy field, and the Minimum Information for a Cellular Assay (MIACA).

The proposed standard is based on XML, and will describe the image and associated metadata, including experiment name, type of container, barcode, dye used, filter, image size, cell type, and other information. "Our strategy is that the community will encourage all vendors to be able to read and write data in the MIAHA standard, so anyone should be able to use any tool with data that is written to the MIAHA standard," says Collins. Thermo Fisher will support the standard in its HCSGateway and HCSExplorer software.

Thermo Fischer plans to send the draft out for review by the end of this year and then present the standard at a scientific meeting next spring. "We'll see what people think, take comments, and hopefully we'll agree on a standard soon after that," says Collins. "It's not easy. But the good thing is that the community is pretty small and the players are committed to standards."

In fact, most companies contacted for this story endorsed standards. "We're supportive and convinced that it's the right thing to go. Otherwise we're just putting up barriers to adoption," says GE's Tesdorpf.

RESEARCH CHALLENGES

Like any field, HCA is moving forward rapidly, and instrument vendors are doing their best to keep up. They hold regular user group meetings and "go back to the lab and scratch our heads" about how to incorporate new requests into their software suite, says Collins.

One key area is bright field imaging, which does not rely on fluorescent agents. These can be toxic or introduce confounding factors in an experiment. That's an especially important issue among researchers studying stem cell differentiation, as they want to avoid any factors that might influence the differentiation process. It can also be used as an extra visualization channel. "Once you get past four fluorescent probes, it's pretty hard to pack any more. Bright field provides an extra channel, if you will," says Collins.

However, it presents a challenge to image analysis. Unlike fluorescent images — which show up starkly against a dark background — bright field illuminates everything. That poses a greater challenge to object recognition and other image analysis routines. Companies are busily adding bright field analysis capabilities to their software packages.

Primary cells will be another challenge. The typical study uses mixtures of cells, "which is pretty arbitrary," says Daffartshofer. The next step is to use primary cells, taken directly from the patient or animal model, which are more biologically relevant than proliferated cell lines. "That (requires) good algorithms for morphology investigation, because those cell

Tools for Therapeutics

Genetix is a U.K.-based company producing HCA systems that can help streamline the drug development process, particularly in the area of biotherapeutic proteins. Genetix's cell-based platforms can be applied anywhere from target discovery through to clinical diagnosis.

"Our systems for cell-based analysis are based on biology, imaging, software, and automation," explains scientific director Julian Burke. He terms biotherapeutic proteins one of the fastest-growing areas within biopharma, particularly in cancer and inflammatory disease.

Genetix ClonePix FL for protein identification and isolation of cell lines.

"We focus on providing systems for identifying cell-surface proteins that may be potential targets for antibodies or other therapeutic proteins," says Burke. "Ariol is a high-throughput automated microscope slide based image-analysis system for the quantification of biomarkers, which allows researchers to examine the differences between diseased and non-diseased cells." When a researcher compares a protein in a diseased tissue section or cell to a control, the corresponding areas of the slides are automatically aligned and presented to the investigator for evaluation.

A new product, ClonePix FL, enhances the identification and isolation of therapeutic protein-producing cells and cell lines, increasing throughput while decreasing time and cost. It allows researchers to identify clonal cell lines producing the appropriate biotherapeutic protein (such as a monoclonal antibody) in a matter of days.

"The power of the instrument is to identify and pick the top few percent of clones that will contain the highest-producing cells and rapidly eliminate the 98% of cells that will never be suitable for scale up," says Burke.

Meanwhile, in the clinic, Ariol is used for disease diagnosis by imaging proteins expressed on cell surfaces such as solid tumors. "Ariol makes the [histopathlogy] process objective by assigning numerical scores to a pathological section," says Burke. The system can subdivide patient populations "according to those who are likely to respond to a particular drug and those who will not," says Burke.

Burke says Genetix' systems form a key part of many organizations' workflows for the identification of cell lines and therapeutic monoclonal antibodies. Customers include Centocor, Genentech, and Wyeth, as well as academic institutions such as Sloan Kettering — **Laurie Sullivan**

Laurie Sullivan is the Editor of CHI's PharmaWeek.

mixtures can be analyzed on morphological differences from cell to cell."

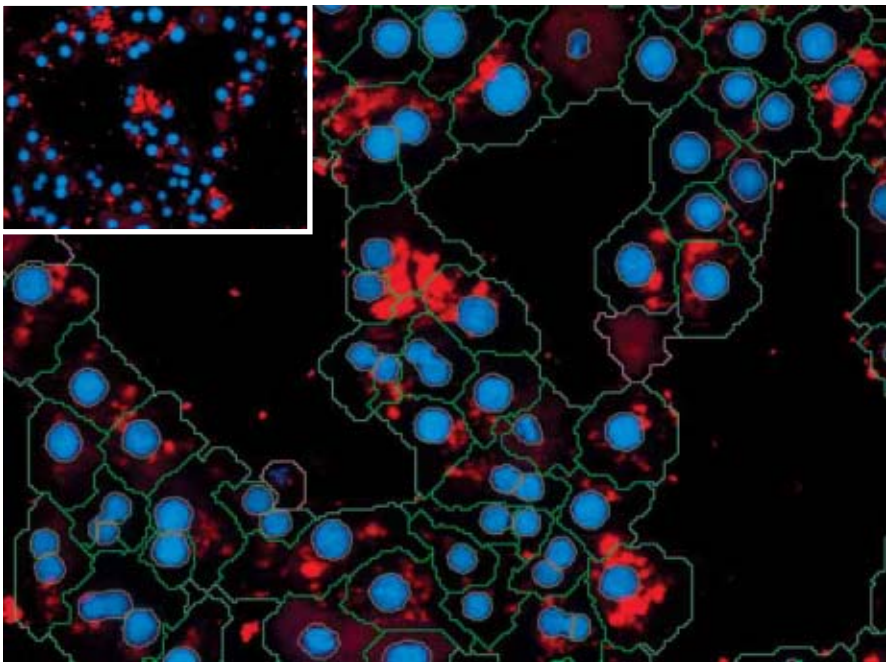
Live cell imaging is also driving software development, as companies work to design software modules that can incorporate kinetics in their analysis. That will lead to a dramatic increase in the amount of data to be stored, as systems collect continuously rather than taking timed snapshots. That represents a data storage and management challenge.

Once the data are stored, users want

easy access, and they also want to be able to mine it. "What I perceive is people wanting to manage all of this data in an integrated way, and to ask questions of the data that were not part of the experimental design," says Scudder. "Some of the system vendors are addressing this, but it's still not a turnkey [solution]."

Tesdorpf agrees: "I think the next big thing is going to be data management, because that's a big problem for users of image analysis. Once you generate terabytes of data, how do you stay on top of that? That's what's occupying us the most right now."

All of these challenges make this a fascinating time for HCA software, as companies pour more resources into their packages. "We did the investment in hardware in the last three to four years," says Hughes. But in the next few years, he says, "I think we'll focus more on the software end of the business rather than hardware improvements." ●



Definiens' Cellenger application plug-in for high content analysis relies on Cognition Network Technology to reveal subtle biological effects and analyze cellular images.

Microsoft's Move Into Life Sciences

Potenzone discusses new plans and the BioIT Alliance.

Rudolph Potenzone has a string of familiar software companies on his resume — the former president and CEO of LION bioscience has also worked for informatics companies such as MDL, Chemical Abstracts Service (CAS), and more recently, Ingenuity, and CambridgeSoft. But his latest employer is the most recognizable of all — Microsoft. Potenzone now has a key position as worldwide industry technology strategist for pharmaceuticals in an expanding group of life scientists at Microsoft. Based at its Redmond, Wash. headquarters, Potenzone helps formulate Microsoft's life science relationships and has taken over the reins of the BioIT Alliance from Don Rule.

Bio•IT World's ***Kevin Davies*** *asked Potenzone to discuss his new responsibilities and appraise the success and future plans for the BioIT Alliance, which since its launch in 2006 has grown to some 60 members.*

Bio•IT World: Rudy, what happened at LION?

Potenzone: I left LION in 2003, when a lot of hard decisions had to be made, and one of them was to really scale back the U.S. operation. I was based in San Diego — the TREGO site that LION had acquired — and the hard decision was made to back out of the U.S. activities, and it was clear it was time for me to leave the company as well. I went to Ingenuity for two years and helped them with the product roadmap, and some of the planning for Ingenuity Pathway Analysis. I was at Chemical Abstracts earlier, where I built the SciFinder product. So delivering these information-based tools is a strong interest of mine, and I appreciated the time with Jake Leschly at Ingenuity. I'm a chemist by training, and this really broadened my thoughts. With LION, we were searching but never got much into the core data structure. Ingenuity was really a lot of fun, looking at ontology and trying to structure that information. That was two years there.

Another of my big interests is electronic lab notebooks (ELN). I tried to build an ELN in 1986 with one of the earlier Accelrys companies called Polygen. We had a product called Centrum — it was way too early. Scientists have so much data information to deal with. The real knowledge of the data is the person who creates it. So I went to CambridgeSoft and managed their some of their ELN development for two years. From there, I moved to Microsoft.

Were you recruited specifically to head the BioIT Alliance?

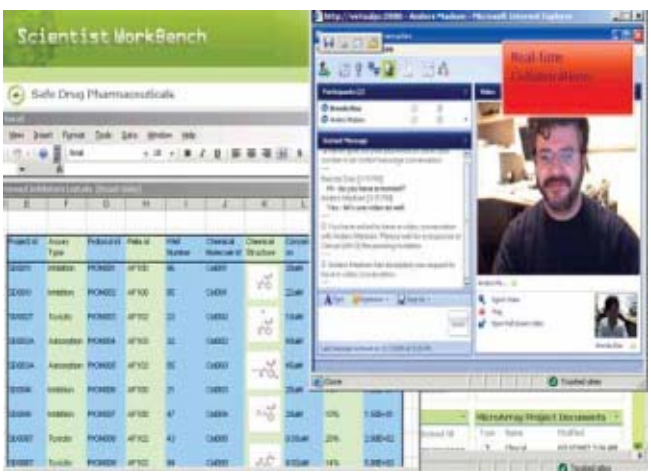
No, that came a bit later. The Alliance is part of my job, but we have an industry solutions unit. I'm the worldwide industry technology strategist for pharmaceuticals. In that role, I'm really a mediator between customers, partners — I spend a lot of time talking to customers building on top of our platforms — and also our product teams, the Excel guys, the Office guys, about the needs of the life science community. While I'm not directly building the products or selling to customers, I am one of the industry experts from the community. The BioIT Alliance really fits into this group, encouraging people to use Microsoft technologies. When [Alliance founder] Don Rule decided to move to a new group, the question was where to put this. He was in an incubation group, one of the evangelism groups here. The BioIT Alliance needed a better home; organizationally, this made a lot more sense.

How big is your group at Microsoft?

It's a growing group. We're the worldwide industry team — there are three of us in the team worldwide at the corporate level, then there are groups in each of the geographic areas. So the U.S. has about eight people dedicated to the life sciences area, now we're growing the worldwide team. The Industry Manager is Rudiger Dorn, who spent ten years at Oracle before joining us, he's in Munich. Zach Hector is just joining our team with over 25 years experience in the health care and life sciences space, most recently with ClinPhone.

Are you hearing from pharma customers about needs to apply Microsoft products?

The product teams solicit help from the field, and it's always difficult to know how to get this data. I'm in a larger group, we've got automotive, aerospace, financial experts etc. We tend to look at requirements in a broader perspective. Suppose you want to put a molecule [picture] into an Excel cell in a spreadsheet — I've been with MDL, I've been with CambridgeSoft, I've been with Accelrys, these companies spend fortunes on trying to make this happen. It's hard to go to the Excel guys and say we want to add this capability, because they're looking at millions of customers and it's not a very exciting possibility. But if we abstract that and say we want



These simulated screen shots show how real time collaboration could work within Microsoft's Scientist WorkBench.

to put a picture of an airplane wing or an automotive tire, then you go and say we need the ability to put more complex objects.

This is the type of thing we do, trying to find what's the real need from the community in the language of the area? I'll bring the life science, then the industry team tries to identify the broader capability. It's a relatively new group. The division is called Enterprise and Partners Group (EPG), it's a real focus on our enterprise customers that Steve Ballmer has brought to the company. Now, as the company evolves, we have to be more focused on the enterprise customers we have. What is their need going forward?

What are some of the other customer needs in life sciences?

One of the big words is integration — integration of information. The product that's really maturing in the Vista environment is SharePoint. It's going to be our flagship for business intelligence applications, integrating information sources, and we really believe this is going to be a transforming technology. It's layered within the Office suite, and a very important part of the whole platform.

SharePoint is the way to pull together information. Business intelligence is one, but we actually have a proof-of-concept called our Scientist's WorkBench. It's not a product, we call it a reference implementation. But it is an example of what you could do with our platform today. It's built on top of SharePoint, Windows Communication Foundation, Windows Presentation Foundation, Workflow Foundation, etc. It can give a scientist a complete view of their project, they could store their literature searches, they could look at whatever the biologists have done on the project, dipping into Oracle or SQL Server databases, looking at chemistry, MDL, or Accelrys, or CambridgeSoft for example... The scientist wants to look at all the data from their project. SharePoint is a way to organize all the information sources. You can build all sorts of things. We've been working with a few pharma companies; it's not actually in use anywhere, but it's something that will be of interest to the community, and will be a topic for the next BioIT Alliance meeting.

Microsoft BioIT Alliance Members

Aberdean
Accelrys
Active Motif
Agilent
Axendia
Axiom Discovery
Biobase
BioDiscovery
Biomatters
Biotique Systems
CambridgeSoft
CeuticalSoft
ChemAxon
Clarabridge
CLC bio
CTIS
Digipede Technologies
Dotmatics
eXludus
FOCUS Biology
Gencom
GenomeQuest
Geospiza
GGA Software Services
GrapeCity
GulfStream Bioinformatics
Hebrew University of Jerusalem
iLink Systems
IMC
Infusion Development
Integromics
InterKnowlogy
Invision
IO Informatics
IUPAC
Microsoft
m-objects
Newtech Global Solutions
NextDocs
nFocus Technologies
Ocimum Biosolutions
Paradigm Infotech
Persistent Systems
PointCross
PointOne
Premier Biosoft
Progeniq
ProSanos
QL2
RND Group
SGI
Strand Life Sciences
Sun Microsystems
Syntel
SyynX
TAKE Solutions
TEMIS
Thermo Fisher Scientific
TranSenda
VAST Scientific
VizX Labs

How many members do you have in the BioIT Alliance?

We're moving close to 60 members. It's been incredible! Every week I have several possible members talking with us about what we're doing, what's the value to them in joining. The question is always, what is Microsoft going to do with this? Why should I join?

Which organizations have joined recently?

IUPAC (the standards body) has joined — it's interesting to them to know what's going on in the community. They've been spearheading this molecular file format for chemical structures called InChi, in coordination with NIST (National Institute for Standards and Technology) that is being used by the National Library of Medicine for their PubChem chemistry portal (analogous to PubMed). Paradigm Infotech just recently joined. A number of companies who are doing some integration tools, the RND R&D group is another one... these are companies working with medical device companies building their integrated software solution.

We're really branching out beyond just bio-IT. I think people tend to think of bio-IT as more focused on bioinformatics. We're definitely seeing a branching out further downstream into the clinical world and into medical devices, which formally is all part of bio-IT, right?

Do you have any plans to change the scope or direction of the Alliance?

I think we have to be a little clearer about what we can accomplish through the alliance. These partners all have their own businesses and products. I think the alliance is probably better focused on communication: Microsoft talking about its technologies and what types of applications either we or our partners think is useful... or if we have partners who want to share something cool that they've done.

The other thing is partners talking to partners, introducing companies they hadn't thought about working together. It's a pretty diverse group of members, and there are connections that could be beneficial.... I've already brokered a few conversations between partners that may or may not turn into something, we'll see how it goes.

There has been express interest in some of the pharmaceutical companies joining as well. We are talking to them about [that]. Virtually all the pharma companies come to Redmond for executive briefings. They wanted me here to be available for those things. They always ask for an update about the Alliance.

What have been some of the Alliance's success stories so far?

We're continuing to publish case studies — the Scripps was one, one was by Affymetrix. CLC bio has put one out on screening the carrot genome. They're on the website [www.bioitalliance.org]. There's a whole set of case studies we've been publishing. We've taken a pretty backseat to highlighting [our involvement]. We're not a funding group. There's a whole separate machinery for partners who want to work with Microsoft.

Has there been a lot of Microsoft senior executive involvement in life sciences?
If you go up a level, we look at this as health care and life science. There is a tremendous interest in the whole health care community. Steve Ballmer announced earlier this year our connected health framework — taking the entire Microsoft suite of capabilities and layering them into a technology framework to support the health care community. In a similar fashion, we're working on a connected life science framework, which does a similar thing. When we talk about the scientists' workbench, the life science framework will be underneath that as a way of delivering this. At the most senior levels of the company, we've got tremendous interest in health care.

> There has been express interest in some of the phama companies joining.

Where is BioIT Alliance founder Don Rule now?
Don has moved into a new group, the Health Solutions Group (HSG), an actual business unit to build products for the health care community. A year ago, they acquired a company called Azyxxi. This is a hospital informatics system that allows medical professionals to pull together all the data on a single patient — X-rays, cardiograms, everything — and it's now a Microsoft-branded product.

This is another information integration tool; you can imagine where that's going to go. There are applications in the life sciences as well. Ultimately these will be actual Microsoft products... So my focus and the Alliance's focus is more building with the partner community on the here-and-now technology. Don's focus and the HSG community's is more what's next. The company's really dedicated to the health care/life science space. There are several hundred people working on the life science effort. That's the reason I joined the company. ●

IT/Workflow

Of Data Silos and Sacred Cows

Delegates at CHI's third annual Bridging Pharma & IT conference share winning informatics strategies that span the drug discovery pipeline.

BY KEVIN DAVIES

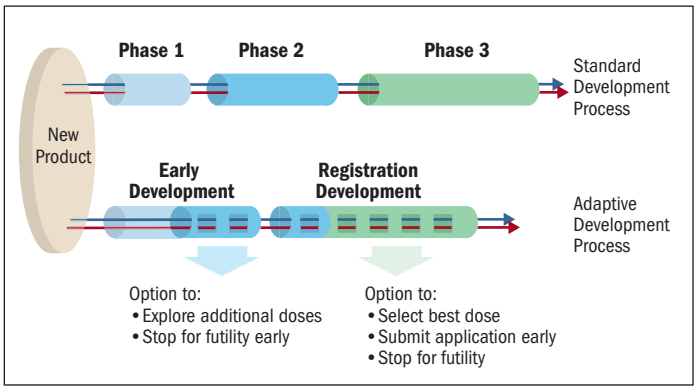
BOSTON — Of the many illuminating examples of the ways that life sciences and IT/Informatics groups and organizations can work effectively together, the problems that faced the Research Portfolio Group (RPG) at Pfizer were perhaps unique. "The [Pfizer] portfolio has become so large it's hard to track information," said **Peter Thadeio**, a project analyst in the recently formed RPG at Pfizer Global R&D in Groton.

But as Thadeio and his colleague Melinda Rottas showed at CHI's 3rd annual Bridging Pharma & IT* conference last month, there is no substitute for ingenuity — even for an organization with a budget as big as Pfizer's.

Thadeio and his colleagues were tasked with producing a system that would afford researchers and executives an overview of all compounds — active and inactive — in development from Phase I-IV, as well as properties such as compound history, structures, status, etc. The challenge was to integrate data from disparate parts of the company. "Discovery and development has been housed in totally separate systems... ne'er the twain shall meet," said Thadeio. "We ran into a huge roadblock — the information was siloed." As Thadeio's group drew up solutions to merge the data, he received discouraging news. "We fell below the line [for funding]," he recalled. Another catch: "Failure to deliver was not an option!"

With no funds, Rottas and colleagues had to cobble together existing software — Business Objects (recently acquired by SAP), Microsoft Excel, Spotfire, and PowerPoint — to create a coherent stream of information that would allow data mining in a standardized format. The key goal was to develop a consolidated file that would contain both discovery and development data, easily updateable on a regular basis.

Step one involved using Business Objects and several data pulls to gather the relevant attributes on specific compounds into a single consolidated view. Next, an Excel data extraction was performed on a proprietary database, and merged with the Business Objects data pull. Excel housed the information where syntax issues that had grown over the years could be resolved (eg. GTN/Groton).



Adaptive trials provide windows into the clinical trial process.

The next step used Spotfire to inspect the data for the presence of duplicates, which could be removed. Finally, the team populated a sequence of tabs with the updated information (from idea to registration) and built a series of formulae around the core data set, which automatically update. These formulae generate a consolidated file of portfolio information.

RPG now produces monthly reports containing graphs and calendars for 12 therapeutic areas, including current status and forecasting, said Rottas. There is interest in a dashboarding system, although licensing rights have been an issue, she said.

"Considering the size of the company and the complexity of the legacy systems, we consider this a 'win' to be able to make this happen," said Thadeio. Rottas' final words of advice? "You don't need to be an IT expert to create these reports," she said. "I'm a chemist by training."

Reagent Records

At Merck, there was less concern about tracking drug candidates. "We track compounds like gold in Fort Knox," said **Vic Uebele**, a research fellow with Merck's neuroscience group at West Point, Penn. But the situation with reagents was not so good, especially plasmids, cell lines, and antibodies. It was like "pennies at the cashier convenience store," he said. Each lab tracked its own reagents, often on paper. Uebele's lab stored information on 2,500 plasmids in 3-ring binders. Problems were compounded by people moving, office relocations, "lost" reagents, legal restrictions, and duplicated projects. In addition to tracking reagents, there was a lengthy list of attributes for each reagent to be recorded, including sequence data, species, source, growth conditions, safety factors, and last but not least, location.

According to Uebele, the impetus for building the reagent tracker actually began with Merck research chief Peter Kim, who said, "You need to talk to Ingrid Akerblom [now Clinical IT] to get this project started." As **Lori Harmon**, manager of drug discovery project support explained, the IT infrastructure was built out in three phases, beginning in 2005 by establishing the requirements for tracking cell lines — initially for the oncology division — moving onto plasmids and antibodies the following year. The only stipulation was that the back-end had to be Oracle. "The front end had to 'feel good'," said Harmon.

The formal bidding process involved three commercial and two internal systems. The final decision was to enhance an internal application, based on workflow and functionality, implementation time

(CONTINUED ON PAGE 36)

* CHI's Bridging Pharma & IT - Sept 30-Oct 2, Boston.

Virtualizing Your Desktop for High Throughput Sequence Search and Analysis

Bioinformaticists, IP professionals, life science researchers and pharmaceutical researchers all want answers pertaining to gene sequences – patent assignees, sequence similarity, expression data, competitive landscape and more. In this webcast, we will demonstrate how you can analyze 1 to 1,000's of sequence queries in a matter of minutes for exploration of biological sequence information while simultaneously investigating the Intellectual Property landscape. GenomeQuest's sequence search and analysis platform combines biological and annotated patent database search, with filtering and results management capabilities from the web interface or command line to transform your desktop into an information gateway. Join GenomeQuest's Senior Director, Content Development, Kamalakar Gulukota as he demonstrates how:

- To quickly apply relevant search strategies, and search multiple sequence patent databases, — at once
- A powerful API enables the running of high-throughput sequence analysis pipelines
- Users can conduct 1 to 1000's of sequence queries against the GQ database
- FilterSmart technology minimizes results for relevant result reports
- Results can be exported and shared
- GenomeQuest offers scalable accessibility based on usage and security needs

www.bio-itworld.com/webcasts/genomequest2

Speakers:

Kamalakar Gulukota, Ph.D.
Senior Director,
Content Development

Kevin Davies, Ph.D.
Editor-in-Chief, *Bio-IT World*
Author, "Cracking the Genome"

Available NOW!

Bio·IT World
OnDemand Webcast

GenomeQuest INC.

Bridging Pharma

(CONTINUED FROM PAGE 34)

and cost, and ease of use and deployment. They opted for a distributed model, because of reluctance on behalf of many groups to part with local freezers.

The result — deployed in July 2007 — is the MRL BioStore. The web application has an intuitive "drill down" to track freezer inventory by racks, boxes, and individual samples. Boxes are ticked to check out vials from any freezer. Nomenclature is a challenge — the application uses both a standard Merck dictionary and a BioStore dictionary. The system, which tracks some 3000 materials, currently has some 500 users across 20 departments, including the new Merck facility in Boston, which uses BioStore to track every cell line.

No Sacred Cows

Founded in 1993, ArQule went public in 1996, building new technology platforms to identify drug candidates with structure-guided drug design. The company has so far pushed three oncology compounds into the clinic.

According to **Mark Ashwell**, VP medicinal chemistry, ArQule has assembled over the years "a comprehensive toolkit of IT solutions for problems that are, on the face of it, often un-addressable for a 100-person biotechnology company." The firm uses a wide variety of third-party tools, including Spotfire and Activity Base from IDBS. But as the company matured, and with everyone wanting a variety of software tools on their desktop, Ashwell said it was imperative for ArQule to carefully review its software needs as the company moved into the clinic while also carefully managing its resources.

ArQule turned to Tessella, a U.K.-based scientific software consultancy, to review its informatics systems. **John Whittle**, technical manager with Tessella, said the first phase was to identify the risks to existing systems, prioritize mitigation strategies, and identify unmet requirements. In short, what could ArQule do better? "No sacred cows, every system has to justify its existence," said Whittle. Areas of priority included the workflow around discovery chemistry and the IT infrastructure, which needed to be brought up to current standards regarding data security and protection. One advantage, Whittle said, was that "conceptually, they don't think in terms of different therapeutic areas." A key priority for Tessella was not to damage ArQule's internal "seamless data structure — don't introduce silos."

Jerald Schindler, VP late stage clinical development statistics at Merck, delivered a superb overview of adaptive trials — the notion of using data unavailable at the launch of the trial. The goal is to maximize information collected on effective drug doses, while minimizing that on non-effective doses (See "Schindler Adapts to New Trials," *Bio•IT World*, June 2007).

Under conventional trial designs, clinicians and statisticians don't learn whether they designed the appropriate trial until the results are decoded. Adaptive trials permit the exploration of additional doses, while the best dose can be selected for phase III. The result is a merger of Phase I/IIa, focusing on safety and dose response, and Phase IIb/III. This should reduce time in the clinic from typically 5-10 years to somewhere between 3-7 years. No wonder the industry is excited.

Although the benefits are "really obvious, everyone should be doing it," Schindler stressed it's not easy. The ideal eCinical system needs two databases, one for data acquisition, the other for review and submission. The goal is to integrate drug supply, randomization, electronic data capture, and IVRS. ●

Tracking Genomics Data

The Harvard Medical School–Partners Healthcare Center for Genetics and Genomics (HPCGG) Laboratory of Moledcular Medicine tests genetic markers for hearing loss, cancer, and cardiovascular problems, among other things. From the patient to electronic medical record (EMR), the genetic data are gathered, processed, directed to a geneticist, then a clinician, and finally used in treatment decision-making. The workflow presented an IT challenge for Sandy Aronson, director of IT at HPCGG.

Through an "IT lens," Aronson broke the workflow down into three components. The first "looks a whole lot like a manufacturing process support," he said. To support the process of gathering samples from patients, processing them, and identifying genetic variants, Aronson and his team developed GIGPAD, or Gateway for Integrated Genomics, Proteomics, Applications, and Data. Like an enterprise LIMS superstructure, GIGPAD manages data from multiple labs through the analysis phase (See "Harvard's Personalized Medicine Gateway," *Bio•IT World*, Aug. 2005).

The second leg of the workflow, starting with raw data arriving at a geneticist, is knowledge management. Aronson and his team developed a combination of two systems. GeneInsight is the HPGCC database to "store correlations established between genetic variants and clinically relevant facts." GVIE, Genetic Variation Interpretation Engine, matches the data gathered via GIGPAD with the information in GeneInsight, and generates a default report for the geneticist. Finally, a geneticist approves or adjusts the GVIE-produced report and enters it into the patient's EMR.

Even with three custom systems, challenges persist. One of the biggest is evolving technologies. "We really want to make sure we can stand up IT support for technologies like Affymetrix microarrays as soon as possible," said Aronson. In addition, the continuously decreasing cost of sequencing, and the corresponding increase in genetic variants identified and used in molecular diagnostics, means an ever-expanding target for researchers and clinicians.

Aronson hopes that pharma may be able to help prepare IT systems for new instruments and tests breaking into the clinical realm, and that collaborations in the future might offer "some assistance with some of our key pain points."

— Allison Proffitt

Challenges in Next-Gen Sequencing and Web 2.0 Storage Solutions

Available NOW!

Bio·IT World
OnDemand Webcast

www.bio-itworld.com/webcasts/bluearc

Featured Speakers:

Dr. James Reaney
Director, Life Sciences
BlueArc Corporation

Dr. Scott Kahn
Chief Information Officer
Illumina, Inc.

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

www.bio-itworld.com/webcasts/bluearc

Next-generation sequencing, and the data that it generates, has presented an incredible data management challenge to the informatics community. Developing a data management plan that addresses both the sheer scale of the data and provides effective management of the data through the informatics pipeline is not easy. While next-generation sequencing holds great promise to make possible a vastly more complete characterization of the human genetic machinery, without a robust storage architecture and efficient data retention strategies few experiments will ultimately succeed. Please join Illumina, Inc. and BlueArc Corporation for a webcast discussion on next-generation sequencing data management strategies.

Participants will learn about:

- The scope of the challenge - Discover how even moderately sized sequencing projects will generate many terabytes of raw and processed sequencing data -- and how larger sequencing efforts will require fundamentally new data retention strategies.
- An analysis of the sequencing informatics pipeline - The speakers will discuss where gains may be made to more effectively manage the data workflow.
- Optimized solutions for informatics architects - See how a high-performance, tiered storage solution can provide future-proof scalability to growing next-generation sequencing research pipelines.

BLUEARC®

illumina®

A Critical Function

The clinical trial help desk is essential to global studies.

BY ANN NEUER

With clinical trial technology now an integral part of many global studies, the full-service Help Desk is emerging as a critical tool for successful implementation. Providers are bolstering their Help Desk operations to support the many administrative and technical questions that arise as end-users at the site level, study monitors, and strategic partners embrace technologies such as electronic data capture (EDC), clinical trial management (CTMS), and clinical data management systems (CDMS).

Rich Deyermond, VP of global customer care for Phase Forward, says, "The Help Desk is absolutely critical to what we do. It's the main touch point with the end-user on a day-to-day basis. They have to have a place to go with questions."

Just a few years ago, when clinical trial technology was less prevalent, Help Desks were far less sophisticated. Dan Piekarz, senior director of support services for Medidata, recalls: "Before 2004, we could handle the volume of calls with a much smaller Help Desk, and project managers were typically brought in early in the process." But as Medidata expanded from being a U.S. company to a global enterprise in some 130 countries, "we built a full-service Help Desk that has grown to a staff of 80 operating 24/7 worldwide."

Medidata has three Help Desks — United States; Sofia, Bulgaria; and Tokyo — and offers six core languages: English, French, Spanish, Italian, German, and Japanese. In all, 20 languages are available among the staff, all of whom are bi- or tri-lingual. Piekarz says that the dispersal of Help Desk locations provides a multitude of languages and better serves the callers. "In Japan, for example, it's very important to address doctors with the proper level of respect. This is reflected in the nuances of the language that would only be known to someone living in Japan or someone who is very familiar with the culture and language," he explains.

Medidata uses a tiered support structure. Tier I handles most calls from investigative sites, and questions about administrative and technical use of Medidata Rave, the company's data collection and management solution, as well as basic troubleshooting. Tier II staff members

MARK GABRENYA

cover applications and how they are configured, so they provide more intense technical support. Tier III may address questions related to integrating the flow of data between Medidata Rave and third-party applications or how to migrate a patient to a new version of the case report form (CRF).

Alternative Help Tools

Phase Forward also staffs its Help Desk 24/7, and operates mostly in the United Kingdom with a staff of 60. Service is offered in six core languages, with Russian, Polish, Portuguese, Mandarin Chinese, and Cantonese also available.

The Help Desk, or Global Support Center, is the initial triage point. Deyermond says typical Tier I questions include, "'I'm having trouble with my log in', 'How do I reset my password?' or 'I'm trying to pull a report and I'm not sure which screen or icon to use.'" Some customers, particularly if they are more experienced, may call directly into the Tier II Corporate Applications Group, and pose questions for the applications or production operations support teams. These could be questions about unexpected pieces of clinical data, and may require an in-depth analysis of the application itself.

According to Deyermond, a full-service Help Desk needs more than well-trained multi-lingual staff — it also needs a growing cadre of online resources. "The industry is moving toward self-service. We've made an investment in automated tools, so over time, we may not need as many people to handle the calls, but you don't want to lose that frontline contact. You need the right mix of human contact and self-service," he explains.

Advanced Clinical Software (ACS), a CTMS provider, has also invested heavily in online resources to help users answer questions and resolve issues. Kris Olson, marketing manager, explains that when customers purchase StudyManager, the company's CTMS product, they select an implementation package that includes various levels of start-up support intended to limit support needs from the beginning. In addition, there is a searchable online resource known as Knowledgebase, and a self-service center accessed at www.studymanageruser.com, which focuses on peer-to-peer interaction.

"Users can respond to questions that other users may have, they can look at best practices and offer input to us on ways we can improve our customer support and service," says Olson. In addition, ACS has created a number of short instructional videos about StudyManager that users can access online.

ACS expects the Knowledgebase resource to grow along with peer-to-peer interaction and the number of videos. Olson aims to increase the number of questions resolved online from 60-70% to 80% in the near future. ●

The Changing Roles for Telemedicine

In addition to patient monitoring, telemedicine adds patient recruitment to the list.

BY DEBORAH BORFITZ

At the University of Maryland Medical Center, the Greenebaum Cancer Center views telemedicine conferences as occasions to remind community-based physicians of clinical trial options and the particulars of the enrollment process, says Mohan Suntha, vice chairman of the center's department of radiation oncology.

The same trial treatment options available at the University of Maryland are available at four affiliated community hospitals where university-employed clinical research associates have been placed, says Suntha. Over the past several years, 25 percent of participants in radiation oncology trials — or roughly 50 patients a year — have enrolled in their own communities rather than at the university's main campus in downtown Baltimore.

"We've integrated telemedicine technology with our commitment to clinical trial access," says Suntha. Telemedicine links provide a means for university-based investigators to communicate with patients while they're in a study. They also allow CRAs to seamlessly transmit data mined at community-based sites.

Other academic institutions — Johns Hopkins and the universities of Arizona, Texas, North Carolina, and Michigan among them — likewise use telemedicine as a means to get clinical data from remote sites, including hospitals abroad.

The same technology can aid in the recruitment and management of subjects in clinical trials. Pharmaceutical companies would have much to gain by encouraging the use of telemedicine at community-based investigative sites and even in subjects' own homes, according to Jay Sanders, founder and CEO of The Global Telemedicine Group (McLean, Va.).

For starters, trial participation would be more convenient and less stressful for people miles from major research institutions if they could accomplish study visits at home or in a familiar, nearby hospital or physician's office, says Sanders. Data collection via telemedicine could also significantly cut monitoring costs of contract research organizations (CROs) and ensure the quality and integrity of the information. "Wireless sensors now exist that allow real-time recording of [all] critical physiological parameters on a continuous basis." From anywhere equipped with a computer, "clinical data can be immediately inputted by email to a central data repository."

This is neither a complicated nor expensive proposition, says Sanders. "There is a telemedicine system today [manufactured by BL Healthcare, Boston] that consists of a small set-top box with a single input and output outlet that immediately converts a channel on a normal television set into a totally interactive, 30-frames-per-second channel."

Long Distance Monitoring

Tele-homecare today is taking place in roughly 30,000 households, most notably for patients of the Veterans Administration. "The more we looked at our ability to provide access where people live and work, the more we found out that the 'exam room' needs to be where the patient is, not where the doctor is," says Sanders. Blood pressure values are generally better at home than at the doctor's office. It's also "more appropriate to assess pulmonary functioning at the location where patients breathe air most of the day."

In terms of the transmission of imaging studies, ECGs, and results of in-home patient monitoring devices, there has been "significant use of telemedicine" in conducting clinical trials, according to Mark Goldberg, president of the perceptive informatics and clinical research services divisions of Parexel, and former president of the American Telemedicine Association. Use of in-home telemedicine technologies, in particular, have the added benefit of making trials more appealing for people to participate in because they are more convenient.

The technology is useful for trials only to the extent that the tools and devices are validated to collect data in a manner consistent with the applicable computer system regulations.

Telemedicine is, to a lesser extent, potentially valuable in the direct recruitment of patients into trials, or at least those seeking patients with rare diseases or conducted in remote locations like sub-Saharan Africa, says Goldberg. But people with unusual conditions may well travel to specialty centers for care, and trials conducted in far-flung regions of the world increasingly receive on-the-ground CRO support as clinical trials have become a truly global activity. On the other hand, telemedicine — particularly in the form of electronic data capture — could help reduce the amount of required on-site monitoring.

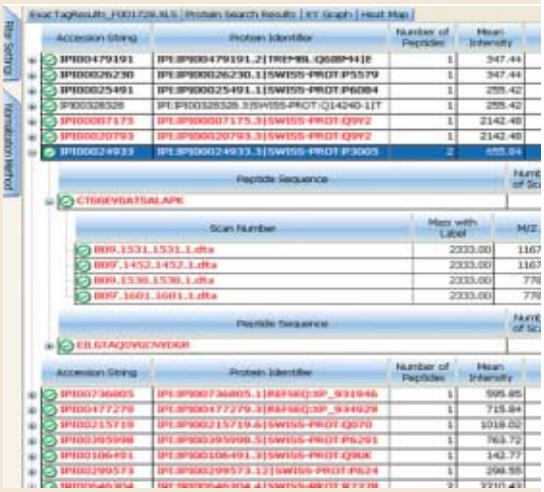
Another area in which telemedicine might be useful is in qualifying people for trials, says Goldberg. If the inclusion criteria for an arthritis trial include a minimum of disease, as evidenced by x-rays of the hands, telemedicine could be used to transmit those images from various, patient-convenient locations to a qualified image reader.

> We find that the "exam room" needs to be where the patient is.

Study sponsors, rather than study sites, are likely to drive telemedicine usage and do so across studies, says Goldberg. Investigators who want to adopt a telemedicine tool — i.e., a scale to monitor weight for a congestive heart failure study — may well have to obtain both sponsor and institutional review board approval before using it in a trial. ●

New Products

Solutions for Superior Quantitative Protein Analysis



ExacTag

ExacTag Analysis Software version 3.0 expands PerkinElmer's ExacTag product line. "[ExacTag] is the highest multiplexing mass tag technology available in the marketplace. You can simultaneously analyze up to ten samples," says Peter Banks, technology and business development manager, Molecular Medicine, PerkinElmer Life and Analytical Sciences. The newly released ExacTag Analysis Software version 3.0 "automates both protein ID and quantification," says Banks. "It's a seamless data acquisition package that we have for our customers that allows them, if they're using [protein search engine] Mascot, to really easily do protein ID and quantification no matter what tandem mass spectrometer they're using."

PerkinElmer also expanded their Phos-tag line of phosphorylation analysis reagents, offering a higher degree of sensitivity and selectivity. **Phos-tag Gold** is a colorimetric assay for the identification and quantitation of phosphoproteins in solution. **Phos-tag Enrich** is for the selective enrichment of phosphoproteins and phosphopeptides in biological samples, such as cell lysates. "It's very easy to use," says Banks of the Phos-tag assays. "This [assay line] is the only one that will work in both a 96-well plate and a microfuge tube."

Products: ExacTag Analysis Software 3.0; Phos-tag Gold; Phos-tag Enrich
Company: PerkinElmer
Available: Now
For More Information: www.perkinelmer.com

At the Limits of Mass Spec

Waters' **Synapt High Definition MS** employs high efficiency ion-mobility based measurements and separations along with specialized software to analyze ions differentiated by size, shape, charge, and mass. The system is intended for researchers working at the limits of conventional mass spec and who need to further characterize and define their samples.

Waters' Synapt HD mass spec.

The **IdentityE High Definition Proteomics System** is Waters' UltraPerformance LCA/High-Bandwidth Mass Spec system that catalogues complex protein digest mixtures. The system features a nanoACQUITY UltraPerformance LC system for high chromatographic peak capacity together with an electrospray tandem mass spec for high dynamic mass resolution and consistent over-sampling of complex protein digests.

Product: Synapt HDMS;IdentityE HD Proteomics System
Company: Waters
Available: Now
For More Information: www.waters.com

Focus on Proteomics

Bruker's **Compass 1.3** is ion trap software offering both electron transfer dissociation (ETD) and proton transfer reaction (PTR). The software also includes several new modes. PAN eliminates the low mass cutoff of ion traps, and ActiveEjection eliminates the most abundant ions to enhance dynamic range. Finally, Compass 1.3 offers dynamic scheduling to allow for monitoring hundreds of target compounds in parallel.

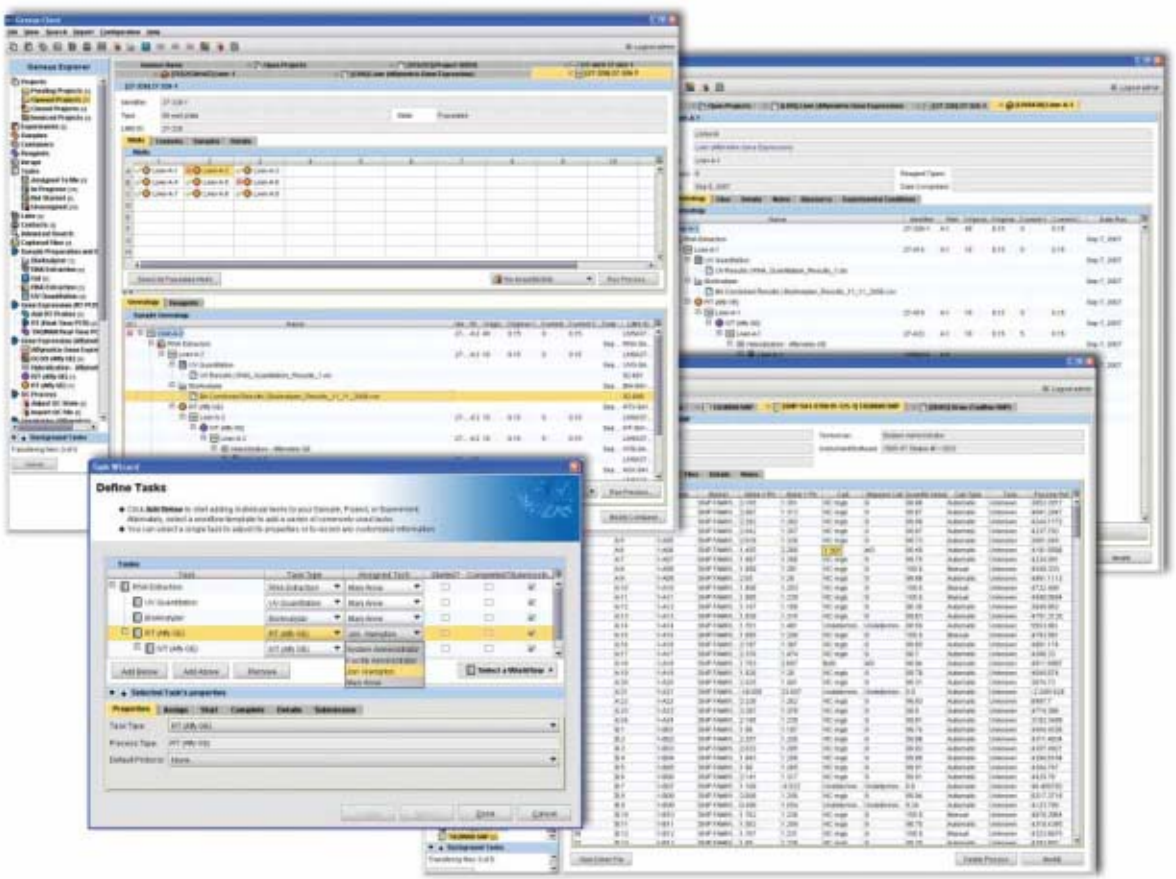
The ultraFlex III MALDI-TOF/TOF mass spec runs both Compass 1.3 and ProteinScape2.

ProteinScape 2 software focuses on LC-MS/MS protein analysis for biomarker profiling, quantification, and validation. ProteinScape 2 supports all current label chemistries including multiplexed labels, as well as label-free quantification. The software has a number of data viewers that permit evaluation and validation on each level of proteomics experiments. The ProteinExtractor algorithm produces non-redundant protein result lists across entire projects, and the processing popeline adopts the HUPO Brain Proteomics Project processing guidelines.

Product: Compass 1.3; ProteinScape 2
Company: Bruker Daltonics
Available: Now
For More Information: www.bdal.com

New PRODUCT Preview

GenoLogics Announces Geneus™ is now a Generally Available Product



- Geneus™ is both an integrating LIMS and a data management solution that interconnects all of the genomics instruments, software and platforms in a lab, streamlining lab operations. It functions like the backbone of the lab, providing lab and experiment management, automation and data archiving, data integration, reporting, and querying, to improve and accelerate lab operations.
- Geneus offers purposed application modules built specifically to enable seamless connectivity to instruments, software and platforms used for common genomics processes including, but not limited to, the Affymetrix and Taqman platforms.
- Geneus is built on a robust enterprise lab informatics platform that is scalable to service many different labs across one organization. This underlying informatics platform provides broad LIMS functionality in addition to all the science specific functionality that Geneus offers
- GenoLogics' informatics platform enables our solution to adapt to changes in lab operations and workflows, new instruments or new applications to meet the current and future needs of our customers.

Geneus is available now at www.genologics.com/products/genomics_solution.php

breakthrough thinking GLS GenoLogics

The Russell Transcript

Improving Structure Prediction

JOHN RUSSELL

Computational biology took a significant step forward recently as a group of researchers led by David Baker developed an *in silico* approach to accurately predicting the 3D structure of small proteins using only their amino acid sequences and NMR data. In fact, their new method was able to improve on some structures previously determined by x-ray ray crystallography.

The work is presented in a new *Nature* paper, "High-resolution structure prediction and the crystallographic phase problem." (Qian, B. *et al.*, *Nature*, 14 October 2007, doi: 10.1038/nature06249).

The effort to develop computational approaches to protein structure prediction has a rich history, and Baker, a University of Washington researcher and Howard Hughes Medical Institute investigator, has long been a prominent figure in those efforts (See, "Computational Biologists Join the Fold," *Bio•IT World*, June 2002).

Over the past decade, much of Baker's work has wound up in Rosetta, a software package to predict protein structure. The code is free to academics and can be licensed by commercial organizations from UWash.

Virtually all computational approaches to protein structure prediction attempt to identify minimum energy configurations. Often though, there are vexing protein segments for which the predicted models have a high degree of variability. Among other things, the new work tackled this problem.

"It's as if you have this complex coil of rope, and there is a section that you think just doesn't behave the way it should," says Baker. "So you just cut it out, reconnect the ends, and computationally explore different conformations of just that section until you have a better model of its behavior."

2002 FILE PHOTO

David Baker is perfecting structure prediction.

Baker says it is now possible to quite accurately predict the shape of smaller proteins. Indeed, they accurately predicted the 3D shape of a protein with 112 amino acids using only its sequence data. The new approach is likely to be quickly adopted by both academic and commercial researchers, says Baker.

Here's an excerpt: "[W]e present a new energy-based rebuilding and refinement method that consistently improves models derived from NMR, from sequence-distant templates, and from de novo folding methods. The final models include high-resolution features not present in the starting models, including the packing of core side chains. Bringing together these results from all-atom structure prediction with state-of-the-art algorithms for molecular replacement and automated rebuilding, we show that distant-template-based and de novo models can reach the accuracy required to solve the X-ray crystallographic phase problem."

CASP's Role

The paper "represents a real breakthrough," wrote structural biologist Eleanor Dodson in a *News & Views* editorial also published online by *Nature*. Dodson writes, "This approach demonstrates real progress in several respects: the use of enormous computational power; the exploitation of known three-dimensional structures; the development of powerful search algorithms that relate those structures to new sequences; and the steadily improving tactics used to determine low-energy conformations of molecules."

Interestingly, much of the actual computation was accomplished using the Rosetta@home project in which more than 150,000 home computer users "donated" compute cycles through the distributed Berkeley Open Infrastructure for Network Computing (BOINC) platform. Not everyone has access to such a grand computing grid but that's not always necessary. "It depends on the problem. A modern computing cluster with tens of nodes would suffice in some cases," says Baker.

Baker credits the biennial CASP (Critical Assessment of Techniques for Protein Structure Prediction) competition to computationally predict the 3-D shape of an array of target proteins from their amino acid sequences for driving the protein structure prediction community. "[CASP] is absolutely critical. It provides an objective blind test of methods, and provides a clear assessment of what the current challenges in the field are," he says.

So what's next? "Good question!" he says wryly. "But predicting the future of structure prediction is even harder than structure prediction itself!" Cute.

THE PERFECT CULTURE FOR GROWING A BIOTECH BUSINESS.

Louisiana is cultivating opportunity in the life science industry with Gulf Opportunity Zone incentives that include 50% bonus depreciation or tax-exempt bonds, and a number of additional tax credits. And with resources such as the world-renowned Pennington Biomedical Research Center in Baton Rouge, it makes perfect sense that more biotech companies are relocating or expanding to Louisiana.

LOUISIANA ECONOMIC DEVELOPMENT

TO LEARN MORE, CALL BOB FUDICKAR AT 225.342.6816 OR VISIT LOUISIANAFORWARD.COM/BIO-ITWORLD

WE CAN. Yoh has been delivering high-impact talent to America's leading companies for the past seven decades – longer than any of our competitors. These customers count on our experience, time-honored processes, responsiveness and market intelligence. And our delivery of talent that is 24.7% better than the competition*.

Focus on your job. And your future.

YohClinical
Know greater talent®
yohclinical.com