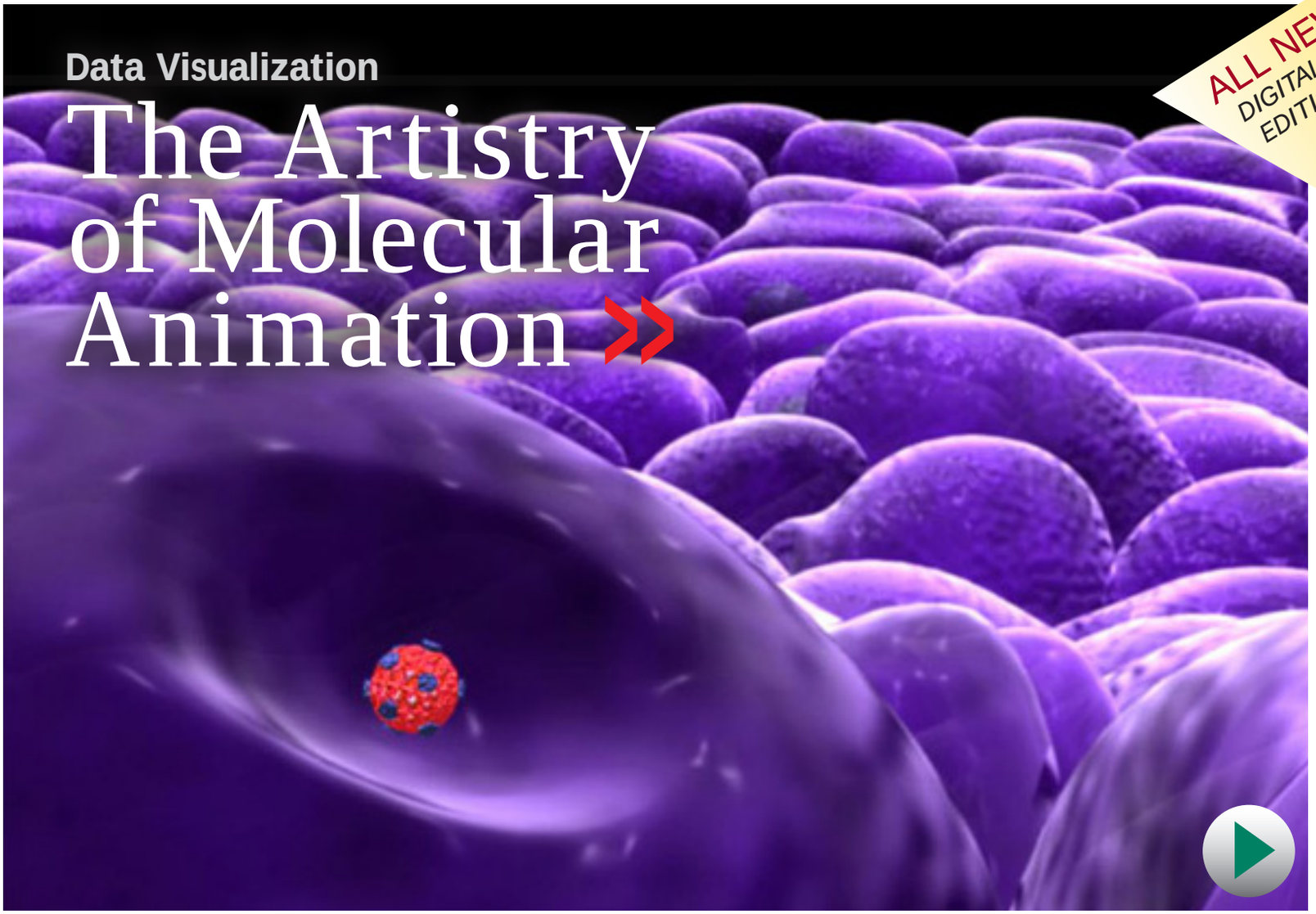
Cambridge Healthtech Media Group

Bio·IT World digital

NEXT-GEN TECHNOLOGY • BIG DATA • PERSONALIZED MEDICINE

JANUARY 2012

ALL NEW DIGITAL EDITION



Data Visualization

The Artistry of Molecular Animation »

FEATURE | Drug Safety

Remedies for Safer Drugs »

First Base

Good Days and Bad in 2011»

eClinical Technologies

Medidata: Integrating Infrastructure for Clinical Trials »

FEATURE

UP FRONT

ARTICLES

KNOWLEDGE CENTER

IN EVERY ISSUE

19th International

Molecular Med TRI-CON 2012

Moscone North Convention Center
San Francisco, CA

February 19-23

February 19-20

NEW-Symposia

Targeting Cancer Stem Cells
Cloud Computing
Display of Difficult Targets
Point-of-Care Diagnostics
Next-Generation Pathology

Pharma-Bio **Partnering Forums**

February 20

Event **Short Courses**

February 21-23

Core **Programs**

- Diagnostics Channel
- Drug Discovery & Development Channel
- Informatics Channel
- Cancer Channel

get inspired

Attend. Discover. Apply.

Spanning five days this year!

New This Year!

TRI-CON ALL ACCESS PACKAGE

Includes 1 Symposium / Partnering Forum,
2 Short Courses, and 1 Core Program.
See website for details.

Organized by Cambridge Healthtech Institute

TriConference.com

Good Days and Bad in 2011

By Kevin Davies

It must be so much fun being a physicist right now. Scientists have discovered two monster black holes—the largest on record—with masses several billion times that of the Sun. And in the Goldilocks Zone of a nearby galaxy, a planet known as Kepler 22b orbits with a balmy atmosphere of 72° F, a prime candidate for life off earth.

At the other end of the matter scale, scientists have caught their first glimpse of the Higgs boson, the so-called "God particle." And we are still waiting with baited breath for confirmation that neutrinos can travel faster than the speed of light, as hinted a few months ago.

In reflecting on the past year in life sciences and IT, however, such notable feats are somewhat harder to find. A list of the top ten science stories of 2011 in the *Guardian*, compiled by veteran science correspondent Robin McKie, includes three of the physical sciences standouts above, but only a couple of items in contemporary biology.

One was the rather sad exit of biotechnology pioneer Geron from the human embryonic stem cell field it helped to create. The other, more cheerily, was the imaging study of a female brain experiencing orgasm, but while imaging is an important technology we cover with enthusiasm, that is, I fear, just slightly beyond our scope.

In 2011, *Bio•IT World* hosted Stephen Wolfram and a record attendance at the Bio-IT World Expo, and convened its first Cloud Computing conference, reflecting the rapidly growing capacity to spin up supercomputing tasks as well as the expanding acceptance of the potential of infrastructure-as-a-service. We also enjoyed covering the potential of new technologies in the form of "wellness chips" (see, "Turning Blood into Gold," *Bio•IT World*, July 2011) and "SNP-doctors," (see, "Powering Preventative Medicine," *Bio•IT World*, September 2011) as well as the creation of the New York Genome Center (see, "Genome Center for Gotham," *Bio•IT World*, November 2011). Some of our New Year's well wishes go to its founders in strengthening what must be at times a tempestuous alliance of Big Apple egos.

Cancer Connections

A story that did not receive as much media attention as it should was the approval by the Food and Drug Administration in August 2011 of a pair of targeted cancer drugs and their companion diagnostics. One was Pfizer's Xalkori, which treats a small percentage of patients with advanced non-small-cell lung cancer (NSCLC) harboring mutations in a gene

called ALK. (A diagnostic test from Abbott Molecular reveals the mutation.) The other was the melanoma drug Zelboraf (Plexxikon), which is accompanied by a BRAF mutation test.

While on the topic of cancer, two papers published this year by the team of Rick Wilson, Elaine Mardis, Timothy Ley and colleagues at Washington University in St Louis, illustrated the potential of individual genome sequencing in revealing molecular aberrations in cancer patients, even if the ensuing targeted treatments don't always prevail or arrive in time.

Ironically, two of the most famous cancer patients to pass away in 2011—Steve Jobs and Christopher Hitchens—both underwent genome sequencing. According to Walter Isaacson's best-selling biography, Jobs had his genome sequenced by researchers from Stanford, Johns Hopkins, Harvard and the Broad Institute. Hitchens, following a recommendation from a person the *Daily Mail* headline called an "evangelical Christian doctor"—perhaps better known as NIH director Francis Collins—was sequenced by the St Louis group in early 2011 as he battled stage IV (metastasized) esophageal cancer.

ANDREW RUSK / CREATIVE COMMONS LICENSE

Christopher Hitchens

Hitchens himself revealed that as a result of the sequencing, he was now taking Gleevec, the Novartis drug for leukemia, begging the question: what did the doctors find in his DNA? While for obvious reasons, Washington University could not comment on Hitchens' prognosis, at a conference last October, Mardis presented an intriguing case study of a patient with a long history of tobacco and alcohol abuse, referred to simply as "ESC1" (esophageal cancer patient #1).

ESC1's tumor, it turned out, carried mutations in a gene called *DDR2*. A quick PubMed search reveals a couple of interesting facts: first, *DDR2* mutations have been associated with several metastatic cancers. Second, the protein encoded by this gene is a potent target of Gleevec and related tyrosine kinase inhibitors. So one can see the rationale for switching this patient to Gleevec. ESC1 responded well at first, but Mardis stressed that her colleagues were not managing the patient's care and did not have up-to-date information.

Hitchens, who said cancer was like a "malignant alien" in his body, died from pneumonia in December at the M.D. Anderson Cancer Center in Houston, writing columns and meeting deadlines to the end. His death hit me harder than Jobs', I suppose because we had a few things in common: scotch-loving British journalists and authors, both graduates of Oxford University. Of course, there the comparison ends.

In one of his last interviews, Hitchens told the BBC: "Whatever day came that the newspapers came out and I wasn't there to read them, I've always thought that would be a bad day—at least for me."

And us, Christopher. Cheers.

H3 Biomedicine: Health, Hope and Heaps of Japanese Funding

[Drug Discovery] With a $200 million commitment from Eisai Co., cancer drug start-up has clear path ahead.

BY KEVIN DAVIES

CAMBRIDGE, Mass.—The 'H3' in the name of the new oncology drug company H3 Biomedicine stands for "Human. Health. Hope."

It might also stand for the heaps of Japanese investment—up to $200 million—that it is receiving from Eisai Co., Japan's fourth largest pharma company. It is an interesting and possibly unique funding model that spares the company the travails of seeking venture finance or short-term licensing deals, says co-founder Stuart Schreiber, professor of organic chemistry at Harvard University and the Broad Institute.

One year after H3 Biomedicine's official launch in December 2010, it formally dedicated its pristine research laboratories in the heart of Kendall Square, Cambridge, with a ribbon-cutting ceremony held in a swanky new restaurant—fittingly named Catalyst—in the same building. Joining president and CEO Markus Warmuth were dignitaries from Massachusetts state and local government, as well as Haruo Naito, president and CEO of Eisai.

"When I first heard of this company's name, H3 Biomedicine, I asked what the abbreviation stands for," said Naito. "I was expecting my name Haruo to be included in that. 'Human, health...' I definitely believed 'Haruo' comes last. But it is 'hope'—a great disappointment!" he joked.

Eisai is not just writing a handsome check but providing ready access to its drug discovery and development resources. Several Eisai chemists have already moved to the U.S. to join H3 Biomedicine's 30 or so current employees in the gleaming new laboratories (see "Lab Life").

H3's philosophy marries expertise in cancer genomics—the forte of co-founder Todd Golub—with discovery-oriented organic synthesis, a field in which his Broad Institute colleagues Stu Schreiber is an authority.

Clean Slate

"This is an exciting time in cancer research," said H3's CEO Markus Warmuth, a German oncologist. "15 years ago, our knowledge was limited to a few genes. Today, we have data on hundreds of cancer genomes, and it is increasing super-exponentially on a daily basis."

"We have a clean slate, we can look at what we learned from cancer genomes, and try to develop drugs for patients from here. That's why I decided to join H3 Biomedicine," Warmuth

Stuart Schreiber, Haruo Naito, Markus Warmuth and Todd Golub toast the launch of H3 Biomedicine.

continued.

"We're trying to be very focused; we're not opportunistic, it's not a random walk in the park," he said. "We're trying to combine new insights into the pathogenesis of cancer, derived from actual patients' cancer genomes, to bring new small molecules to patients in need... We have to be very patient to be successful in developing drugs."

Warmuth said that H3's success would be greatly aided by having "investor pressures" off its back, "so we can focus on real scientific questions and move forward developing drugs," he said. "We're very happy to have the investor [Eisai] in the background, who has given us a tremendous amount of money to really alleviate the pressure and enable us to take a long-term [approach] at a time that many other companies are trimming down their research budgets."

Warmuth noted that the company's two co-founders conveniently had their offices at the Broad Institute, just two minutes walk from H3's location. Schreiber is a renowned chemical biologist who is the co-founder of Vertex, Ariad, and Infinity Pharmaceuticals. Todd Golub is the founding associate director of the Broad Institute and also the co-founder of Foundation Medicine, among others.

H3 will be tackling cancer in general, not focusing on any specific therapeutic areas. "We're looking at the cancer genome, no matter if its lung cancer or breast cancer or liver cancer. We're looking for the best opportunities—

« we're not trying to just make lung cancer drugs or breast cancer drugs. Let's look at what the cancer genome actually tells us."

Eisai CEO Haruo Naito noted that his company has long-standing roots in the Boston area, having established the Eisai Research Institute in 1988 in Andover, Mass. Two of the four founding scientists are still present, under the leadership of Yoshito Kishi, a Harvard chemistry professor and colleague of Schreiber's.

The pharmaceutical prowess of Kishi and colleagues was demonstrated in the synthesis of Halaven, approved by FDA in late 2010 for late-stage metastatic breast cancer. "It is the first time a drug is proven to extend life of such late-stage cancer patients," said Naito.

The development of Halaven took 16 years from early discovery. The compound is found

Lab Life

Construction of the first phase of H3's lab space was completed last July, with 48,000 square feet currently, including spotless labs for synthetic chemistry, bacterial, and tissue culture, and genome analysis. The design offers a hybrid between open office space with glass frontage and respect for scientists' privacy. "I think it's important that sometimes you can shut the door and think about a problem," said Warmuth.

The common room/coffee break area is surrounded by four modern seminar rooms, each named hope in a different language: Esperanza, Amal, Hope and Kibou. A planned expansion will include a brainstorming room with no projection capabilities, just wall-to-wall white boards and music.

"It's absolutely terrific to work in an environment like this," said chemist Dominic Reynolds, who trained at Cambridge University with Steven Ley. The fume hoods are surprisingly quiet and split into two, so maximizing safety while minimizing energy expenditures. Evaporation equipment is housed in ventilation hoods, with circulating ethylene glycol to cool reactions.

Every chemist is equipped with a lab laptop providing access to electronic literature and Eisai databases and resources. They also have access to a walk-up mass spectrometer and a 400-MHz Bruker NMR machine in a specially vaulted room. "It's really a privilege to have a piece of equipment like this," said Reynolds.

"We want to take advantage of modern synthetic methods to generate proprietary chemical libraries that take us into chemical space not occupied by anyone else," said Reynolds. The emphasis is on quality, not quantity, with the goal of generating chemical libraries containing a relatively modest 15,000 or so novel compounds a year.

"We're embracing old combinatorial methods, but we're not in an Arqule-style,

in an "ugly black sea sponge" that lives off the coast near Tokyo. Scientists collected 600 kg of the sponge to produce 1 mg of anti-tumor agent. Kishi's group synthesized the highly complicated structure in a tour-de-force 61-step synthesis (accommodating 19 chiral carbon atoms that resulted in more than 500,000 potential compounds).

"I hope, Markus, in your case, it won't take 16 years!" joked Naito. "'H3' means three years!" The most important ingredient in H3's ultimate success, he said, would be "the fighting challenging spirit to do something new."

"We want to be very fast and focused," Warmuth concluded. "I'd be perfectly OK if we only ever work on three projects, if they are all successful and produce good drugs."

300,000-compounds-a-year mode," he said. "We're looking for privileged novel scaffolds that we can decorate and access and screen quickly."

One of the key differentiators of the new company, Reynolds explained, is that H3 is performing chemical syntheses in solution phase, in contrast to programs at say Infinty Pharmaceuticals (another biopharma that has featured Schreiber's diversity-oriented synthesis expertise) or the Broad Institute, which run polymer-supported chemistry to facilitate the generation of massive numbers of compounds. "Here, the chemistry is much more tailored to each core, hence the smaller numbers... so it's all solution based and HPLC-purified at the end, which is really important for the medicinal chemists," said Reynolds.

Liverpudlian Peter Smith, previously at the Dana Farber and Millennium, said he had a blank slate to buy the right equipment "to prosecute the biology side of the project." Assays will include protein levels and phosphorylation, DNA manipulation and transformation, and tissue culture—creating cell lines with different genetic backgrounds to help understand which patients will respond to various drugs.

"We'll take different levels of genetic information and mutation status, and we'll look for fusions of genes and epigenetic changes. We're not just focusing on one particular aberration but building the importance of those genetic aberrations to find the right target," said Smith.

Smith said the goal is essentially to define the experiment on day one, to understand what needs to be done in the patient and replicate that through enzymology, cell biology, animal models, and ultimately into the patient. "Knowing the question at the outset helps us prosecute efficiently, so we should get the answer quickly when we go into clinical studies," he said.

Picture This: Molecular Maya Puts Life in Life Science Animations

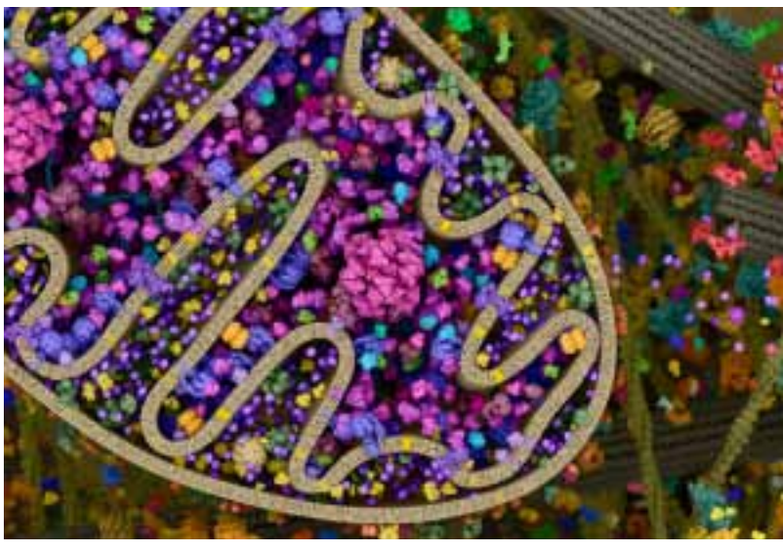
[Data Visualization] Based on the Autodesk platform, Digizyme plug-in proves aesthetic and educational effectiveness.

BY KEVIN DAVIES

In 2010, a reporter sat in a Life Technologies hotel suite admiring a promotional video illustrating one of the company's latest research projects—a single-molecule sequencing system featuring enzymes tethered to fluorescing quantum dots. The video was impressive not merely for pushing the boundaries of sequencing technology, but equally for showcasing some powerful production qualities in 3D animation and rendering that, until recently, would have seemed the provenance of a Pixar movie.

That video was produced by a Boston company called Digizyme (www.digizyme.com), founded in 1999 by Gaël McGill as a side business while he worked on his Ph.D. at Harvard Medical School. Since then, Digizyme has become a respected source for educational and promotional animations, simulations and videos for many diverse life science applications.

"The visual communication of science—that's where my passion is," says McGill. He recognized early on that graphic design, animation and visualization tools were needed by several industries, not least his fellow scientists. Clients include academics as well as biotech, pharma, medical device companies, along with science museums, and public television stations.



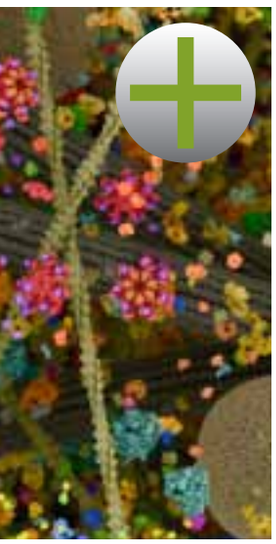
A mitochondrial molecular landscape depicting signaling proteins in apoptosis (created by Digizyme for Cell Signaling Technology)."

After a hiatus to run the company full time after his postdoctoral fellowship, McGill joined the Harvard Medical School faculty five years ago in the Center for Molecular and Cellular Dynamics, where his research focuses on how visualization impacts research and education. He is also creating a new graduate program to train students in biovisualization approaches.

Most companies offering animation tools are staffed by visual specialists or animators with a strong artistic background collaborating

with scientists. "There are some great companies in this space, but what sets us apart is that we're scientists," says McGill. "Everyone at Digizyme is dually-trained as a scientist-animator, scientist-artist, or scientist-programmer. As a result, our visualizations are directly informed by the primary literature and based on raw data."

At Digizyme, McGill recruits twin-threat staff with a science graduate degree as well as strong artistic talent. The company has created stereoscopic movies for science museums, custom projects for pharma clients including Amgen, J&J, Novartis, Life Technologies, and Genentech, as well as completed work for several major hospitals, research institutes, and universities. McGill is also the Digital Media Director and a collaborating author for E.O. Wilson's 'Life on Earth' digital biology textbook.



Going Hollywood

The software that powers the Digizyme platform is Maya, a suite developed by Autodesk, which is a modeling, animation, simulation, and rendering application widely used in media and entertainment circles to create video game characters and environments and special effects for feature films. Of course, such software was never intended to display life sciences data.

"What are we doing with Hollywood software?" says McGill. "There's been billions of dollars invested in films to create powerful software suites, such as Maya."

McGill says that current bioscience research features "an incredible richness of datasets that is just overwhelming. We see visualization as the key to knowledge integration—it offers a powerful and flexible platform to import and synthesize data from various fields—data that would not otherwise interact."

One example would be to create an intracellular landscape using cryoEM tomography data to map the location of proteins in a cell. This map could be populated with atomic resolution protein structures, and then set in motion and simulated using various algorithms including Brownian dynamics. "This begs the question: What software do I use for this sort of thing? It doesn't currently exist, but it's something that interests many scientists. Our approach is to modify and adapt existing animation tools from Hollywood—platforms like Autodesk Maya—to advance toward that goal."

McGill's team built a new layer of code to make Autodesk's Maya more biologically relevant, resulting in a plug-in toolkit called Molecular Maya (mMaya). (see, "A Better World")

Molecular Maya focuses the modeling, animation, rendering, and simulation capabilities of Maya in the context of biological animations and structures. For example, instead of modeling DNA from scratch, it provides a direct link to a scientific database that downloads data directly into Maya and automatically creates the 3D model. McGill and his team are also developing tools in mMaya to rig macromolecules to facilitate biophysically-accurate animations as well as to build molecular environments.

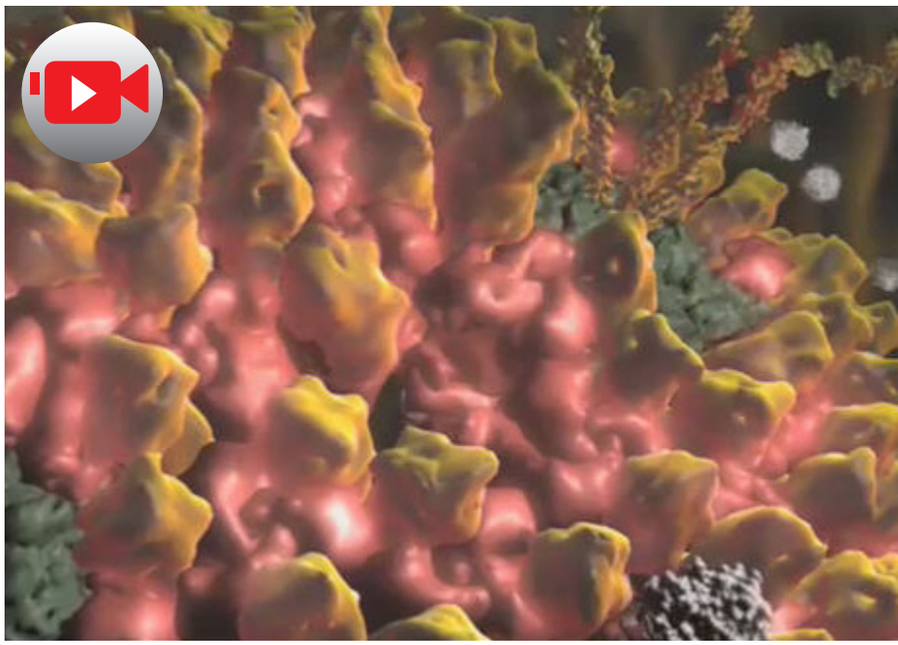
Molecular Maya is written in languages (including Python and MEL—Maya Embedded Language) that allow other researchers to extend its capabilities. "We've created a series of free, open-source scripts that lets anybody type in a protein ID, say, and with a click, have this directly connect to the Protein Data Bank so they can model, manipulate and animate a protein within seconds," says McGill.

This accessibility came to light in the recent IGEM (International Genetically Engineered Machine) world jamboree, where student teams from around the world used synthetic biology principles to design novel cellular circuits and hybrid proteins. "For a student team trying to design something in a forward engineering approach, the design piece is not just making a pretty picture, it can be an integral part of the conceptualization and experimental approach," says McGill. For the past few years, he has been involved with this competition by hosting seminars and offering workshops to help students use Maya and Molecular Maya. (Autodesk has also sponsored IGEM the past three years.)

Not Just Pretty Pictures

McGill underscores the critical role that visualizations can play in experimental research. "I've worked with colleagues at Harvard Medical School where the process of creating a visualization can change one's understanding of the protein," says McGill.

In one project carried out in collaboration with renowned structural biologist Stephen Harrison, who specializes in proteins on the surface of HIV, McGill created an animation that depicts how a viral surface protein—gp41—spears the target cell membrane and refolds on itself to drive membrane fusion and facilitate viral infection. Proteins aren't static, they're shape shifters," says McGill. "gp41's conformational change cannot be depicted with the typical linear interpolation morphing techniques. It undergoes a helical transition and refolds completely." And trying to communicate such a critical and unusual conformational change in a standard research paper was next to impossible.



Digizyme models the cellular entry of a reovirus.

McGill's group helped Harrison visualize gp41's unique conformational change and, in the process, communicate the mechanism in a way that hopefully facilitates drug discovery. "Understanding the specific mechanism of how gp41 refolds may drive better approaches to drug

A Better World

The life sciences community might appear a little staid next to some of Autodesk's other marquee clients, including James Cameron, Frank Gehry, Disney, and Boeing. Not so, insists Patrick Byrne, director of business development for the Californian company that celebrates its 30th anniversary in 2012.

"Our mission is helping customers imagine, design and create a better world," says Byrne. The principal customers are in architecture/engineering, manufacturing and media/entertainment. Within the latter group, Autodesk has been building a product development group to extend Maya into life sciences. "It sits in our media and entertainment team, because Maya was developed for media professionals," says Byrne.

But why life sciences? "When we look at the range of industries we're in, the concept of design workflows is transparent across all industries. Whether you're in one of those industries, our tools enable you to be more efficient, more iterative and more creative. And when we look at academic research, we see an interesting place," says Byrne. "Without much ado from us, people like Gael have taken Maya and used it as a platform to develop visualization tools for their industries. We want to extend and enable that."

One of the keys to Maya's popularity is not just the modeling, texturing and rendering capability, but its extensible nature. Users can write their own plug in, which can be complex or very simple. For students and academic users, Autodesk offers free access to the software for three years in a termed-free license (http://students.autodesk.com/). "As you're perfecting your skill, let's make sure you have access to the tools to perfect your craft," says Byrne.

Byrne sees three prime opportunities for Maya in the life sciences: first, as a platform for innovative tools; second, use in science education curricula; and third, helping scientists in industry visualize and communicate research to a wider marketplace.

There is some strong competition in the marketplace, including an open-source product called Blender, and its derivative Bio-Blender, but Byrne sees the market expanding.

Support of the IGEM competition, for example, puts Maya in the hands of a global student body eager to harness the software in various synthetic biology applications. Byrne says he is proud of the MIT group, which won the 2011 IGEM prize in the health/medicine category, enabling cells to autonomously regenerate.

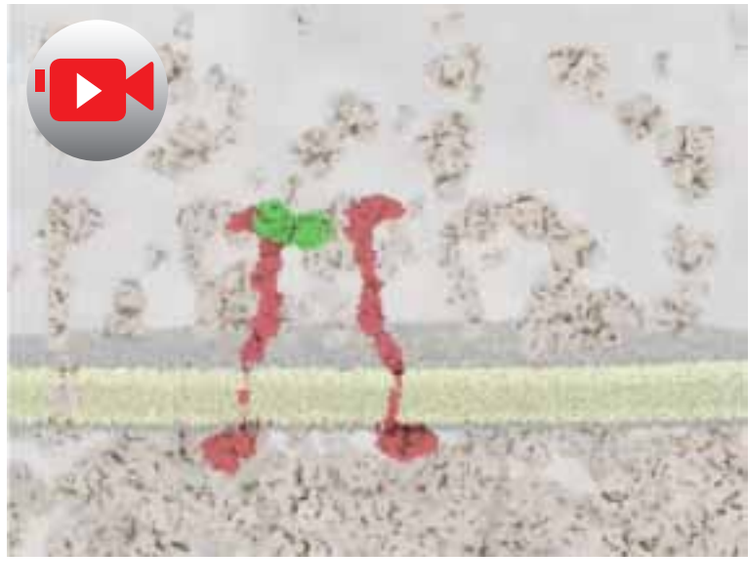
discovery or vaccine development," he says.

Another example is Shawn Douglas (Wyss Center, Harvard Medical School), a former visualization student of McGill's and now a leading figure in the field of DNA Origami—designing novel DNA molecules to fold in vitro to create biologically useful entities. Douglas' Maya plug-in, Cadnano (http://cadnano.org), greatly simplifies DNA origami projects and brings the power and ease of use of typical modeling tools in Maya to this rapidly growing field.

"We're now collaborating to have both plug-ins—mMaya and Cadnano—work together inside Maya," says McGill. "Cadnano figures out how to create the genetic sequences that, when mixed in a test tube, fold into the shape of an object modeled in Maya. It's magic—a great example of how a design tool can transform a process that previously took weeks or months to navigate."

Visualization research isn't just about making movies but also about understanding how to design the most effective movie to impact a target audience. Says McGill: "We're very interested in understanding how design choices—how we represent or animate a process—can have an impact on the effectiveness of a movie, especially in the context of educational animations. We strive to create animations and interactives that have a high pedagogical impact on students, not just animations that are engaging and aesthetically pleasing."

In one recent project, hundreds of students assessed four different styles of animating a simple molecular interaction: a hormone binding to a cell surface receptor. McGill and his collaborators tested students before and after viewing these animations as well as monitored their reactions using eye-tracking methods. "Typical medical animations often impart decision-making properties to molecular entities—as if molecules know where they are going. This couldn't be further from the truth!" says McGill. He acknowledges that creating animations that depict both the stochastic nature of the molecular world while keeping students' attention on the mechanism and specificity of an interaction is quite a design and technical challenge.



3D visualization techniques prove a valuable tool for teaching molecular biology.

McGill has created an educational web site called molecularmovies.com. It has three sections: one is a Showcase of some of the best online cell and molecular animations, organized by scientific topic. A Learning section contains McGill's teaching curriculum at Harvard Medical School, available free. And third is the Toolkit section, where researchers can download Molecular Maya.

Enhancing Productivity of Internal and Externalized Research through Integrated Cheminformatics Systems

Drug Discovery Research and Externalization

For most pharmaceutical and biotech companies, discovery research has become an externalized process across networks of in-house, academic and research institute collaborators and CROs where synthesis, assay development and execution, safety assessment and other aspects of the process may all be conducted by different organizations.

Recently there has been a change from one-to-one relationships between a pharma/biotech and a CRO to dynamic networks in which several partners collaborate on a single discovery project,and are added or removedaccording to specific skills. Whilst early collaborations focused on cost-reduction with fee-for-service relationships, now partners are sought for advanced technology and expertise, and IP is discovered jointly.

Challenges for informatics Systems

These new arrangements bring significant data management challenges:

- Each participant needs to share their data with the commissioning company and appropriate partners
- Each partner needs real-time access to the data from other partners that they require to complete their work
- Dynamic collaborations must be spun up and down as fast as the collaboration network changes

IT departments struggle to support the partnerships with in-house data management systems not designed for external arrangements. They were designed for scientists who were employees with mostly free to access the data and for IP that was all owned within the company, but in collaboration networks, each partner is likely authorized to only see part of the data, and the IP ownership may be with the commissioning company or across the partners.

How Do Companies Solve the Problems Today?

Most companies today solve the problem in one of two ways, each of which has significant drawbacks.

1. Exchange of files. Data is traded in files over email, via FTP or in E-rooms which
   - has the potential to be insecure
   - makes it hard to keep track of IP ownership as files move around
   - is not real time
   - can be error prone or lead to ambiguity in interpretation
   - is costly and time consuming
   - is impossibly complex as the number of collaborators on a single project grows

2. CRO/Collaborator access to internal systems. The lead company opens up access to their internal systems allowing the CRO to access or deposit data directly into those systems. Most in-house systems provide row-level security to allow data access to be controlled but were designed with the assumption that most data can be shared and that those accessing the IP are company employees. There is a risk involved in opening up these systems and an IT cost to ensure that the systems are secured for external access and monitored

HEOS – A SaaS Based Solution Design to Facilitate Externalized Discovery Research

HEOS is a SaaS based system designed from the ground up to address the data management issues within externalized drug discovery projects, securing the IP of each party and providing real time sharing and access to project data to maximize the projects efficiency and drive forward innovation. By using HEOS, project teams across all organizations can have instant access to information that they are authorized to see about their projectsand securely share the data that they are responsible for generating with the rest of the team. HEOS provides upload, storage and search of all types of complex data critical for drug discovery projects including biology, chemistry, safety and pharmacokinetics for example.

For the organization commissioning the research it has the benefit that this collaboration happens in the cloud and there is a clear delineation between their collaboration system and their in-house systems and data. As a SaaS based solution it is easy for a new collaboration to be spun up quickly and then shut-off just as quickly at the end of the engagement with the data from that partner secured in the system.

To learn more join the web symposium on January 19th:

Enhancing Productivity of Internal and Externalized Research through Integrated Cheminformatics Systems

accelrys®

Assembly Required: Assemblathon I Tackles Complex Genomes

[Genome Informatics] The genome assembly competition challenged entrants to assemble a synthetic, 'human-ish' genome.

BY KEVIN DAVIES

From familiar but repeat-laden plant species to obscure vertebrates, more and more genomes are being sequenced that require *de novo* assembly without alignment to a reference sequence. "Every genome has its own story in terms of repeats," says Ian Korf, associate director of bioinformatics at the University of California Davis Genome Center.

Korf is one of the principal organizers of a genome assembly challenge known as the Assemblathon—a competition to identify best practices in the *de novo* assembly of complex plant and animal genomes. Results of the first phase of the Assemblathon were recently published in *Genome Research*.

Korf discussed some of the Assemblathon results and, more broadly, the inherent challenges in genome assembly in a recent webinar hosted by the community forum NGS Leaders. "Sequence analysis starts after genome assembly—you can't do much beforehand... Every genome is a complex genome—even the simpler ones are pretty complex. There's no easy genome," says Korf.

The Assemblathon grew out of the G10K project, which is an effort to sequence 10,000

vertebrate genomes. Clearly, in order to sequence and assemble 10,000 genomes, it is crucial to know what is the best sequencing and assembly technology for the money. "It definitely becomes a cost-benefit ratio looking at 10,000 genomes," says Korf.

Joseph DeRisi (Berkeley), David Haussler (UCSC), and Illumina helped launch the Assemblathon idea. The original goal was to make

two targets: one was a real genome (snake), the other a synthetic genome, to enable participants to determine how well they performed. In the event, the snake data weren't ready, so Assemblathon I, which took place in early 2011, utilized just the synthetic genome.

To create the synthetic genome, the organizers took a copy of human chromosome 13, and artificially evolved the sequence using Evolver

Ian Korf says assembly can't be a stepwise process; it starts as far upstream as library preparation.

software, which introduced mutations in different regions (coding/non-coding) and at different rates. "The sequences were human-ish, but after 200 million years of evolution, didn't look that human," says Korf.

The Assemblathon participants—17 groups in all—were then challenged to put the synthetic reads together. "Because we knew the answer, we could evaluate each one of the assemblers," says Korf.

Results Are In

Commenting on the results, Korf said: "A lot did a pretty good job, but it's more difficult to assemble regions with more mutations, so the coding regions were assembled better than non-coding regions." (The contest did not test the growing number of commercial assembly packages, from the likes of CLC bio, DNAStar, Gene Codes and others.)

The assemblies were ranked by various criteria, including contig and scaffold paths, structural and copy number errors, and so on. In the final rankings, the top five were:

- Broad Institute (ALLPATHS-LG)
- BGI (SOAPdenovo)
- Wellcome Trust Sanger Institute (SGA)
- DOE Joint Genome Institute (Meraculous)
- Cold Spring Harbor Lab (Quake, Celera, Bambus2)

Several useful tools emerged, says Korf, but experience in using the tools makes a big difference. "We found that sometimes two groups will use the same assembler, but the group that knows a bit more about the assembler might do a slightly better job. It's something of an art at this point," said Korf.

Korf says that wisely choosing the many different parameters involved in *de novo* genome assembly is difficult and "probably shouldn't be attempted by amateurs." He advises inexperienced users to "contact one of the major sequencing centers and get them to help you.

Reads and Errors

These two paragraphs illustrate the different type of reading errors associated with Illumina and PacBio data, according to UC Davis' Ian Korf. Illumina sequencing data are made up of very short reads with few errors. Pacific Biosciences technology has a higher (10-15%) error rate, but the sentences (reads) are long enough that they can largely be interpreted.

ILLUMINA:

e pro opria e acc plays erial ng. T in so ks is ll in case meter lete s is gene a cr s tas ining plete the s s are ay, s g the nome gene a gen f amo most geno uracy ainin prope s cou inder es so dant. mate the f and t ckly, st im can iven ole b ous t ckle more and m ing a ene s mong es ar any f e of rmini proj he fi s rel leav on pl ar we ng ma ly. T hout st im ather ortan genes a la , and find firs n pro appea . Giv equen ts pr be e ical ermin e ann ably. ensiv plete lete. raini le be f its es ar ckly, one meant pends t imp ture ore t appe task. se it assem apid

PACBIO:

In the past, genome prbjects moved rather slowly, and the tricqle of data meant lhay parameters could be estimjqeg before tfs genome wns coq that parameterg iould be qstimgted before the genome was completej Today, saqwencing and assemblm are so ripid that genomes cag appear well in advazce cy nh a gene finder depebds on many facrors. Chief agong these is proper graining. Training a gene finder can ie a laborbous task. In the past, genome

Doing it on your own is pretty much guaranteed to give you a sub-optimal assembly... Don't jump into genome assembly thinking it's just like any other bioinformatics problem you can hack with some Perl scripts."

It starts as far upstream as DNA library preparation. "You don't want to choose the assembler as the last thing you do," says Korf. "It must be in conjunction with the sequencing technology, how are the libraries made, the full equation. You can't do it stepwise."

Library preparation is a non-trivial step. "It's really garbage in, garbage out," says Korf. "So much is dependent on having high quality sequence and making your libraries correctly." Indeed, some Assemblathon participants believe there should be a library construction competition, because that's more important in some ways.

Another wise move, says Korf, is to perform a pilot project "to explore what your genome looks like" and gauge the overall repeat content of the genome in question. For example, in a recent project to sequence the gigantic

DNA | SmallRNA | RNA | ChIP | Methylation

The Most Comprehensive Start-to-Finish

NGS and Microarray Analysis Solution Available

Partek Flow™

Partek Genomics Suite™

Partek Pathway™

Alignment → Quantification → Quality Control → Statistics Visualization → Integrated Genomics → Biological Interpretation

Partek®
turning data into discovery®

www.partek.com

Most Creative Algorithm
iDEA Challenge 2011
Commercial

pine genome (22 billion bases, three times larger than the human genome), his group was surprised by the extent to which the DNA repeats were diverged. The most common repeat only made up 3% of the genome, making it easier than expected to assemble. "You should do a little homework ahead of time to get an idea of GC content and other factors," says Korf.

> "It's really garbage in, garbage out. So much is dependent on having high quality sequence and making your libraries correctly."
>
> **—IAN KORF,** *University of California Davis Genome Center*

Longer Lengths

The availability of longer read lengths, such as those produced by the Pacific Biosciences platform, should complement existing short read systems and prove a boost for genome assemblies. "The long reads are fantastic, but the error rate is a bit of an issue," said Mario Caccamo, head of bioinformatics at The Genome Analysis Centre in the UK and a co-author of the Assemblathon I report.

But Korf says the PacBio reads can prove very useful in integrating with short read data: "Genome assembly with longer reads will get much, much easier. The game will be completely changed with reads on the order of 10 kilobases."

As an analogy, he offered two paragraphs, representing the shorter Illumina reads and the PacBio read lengths (see, "Reads and Errors"), pointing out that despite the errors in the PacBio segment, it was still possible to interpret the text.

Korf says UC Davis was one of the first centers to receive a commercial PacBio machine, although it is currently being used for studies of genome biology (centromere structure, fragile X repeats) than assembly. "There are certain things that PacBio can do that nobody else can do. We like those things! Yes, there are errors, but better to ask the question with a few errors and have to do a little harder analysis than not be able to do it at all."

Korf believes the NGS community—"super smart people, full of competitive spirit"—will figure out how to use these 3rd-generation technologies. "Right now, they haven't had enough time to figure out how to put it all together, but they will pretty soon," he says. "What you'll get three years from now will be a lot better than today."

For the second round of the Assemblathon, which concluded at the end of 2011, participants have belatedly tackled the snake genome, along with the bird and fish. "The disadvantage is that we don't know the answer in the end," said Korf. Those results will be announced later in 2012.

FURTHER READING: Earl, D.A. *et al.* "Assemblathon I: A competitive assessment of *de novo* short read assembly methods." *Genome Research* 21, 2224-2241 (2011)

The NGS LEADERS webinar, "*De Novo* Assembly of Complex Plant and Animal Genomes," featuring Mario Caccamo (TGAC), Ian Korf (UC Davis) and Jeffrey Rosenfeld (UMDNJ), is available on demand.

10th Annual BioPartnering North America™

Growing Life Science Business in the Pacific Rim

Partnering & Dealmaking 10 YEARS

› www.techvision.com/bpn

BioPartnering NORTH AMERICA™

10th Annual BioPartnering North America™
Vancouver Convention Centre | Vancouver, BC, Canada
26-28 February 2012

Now in its 10th year of success, BioPartnering North America (BPN) will build on strong relationships focusing more clearly on the Pacific Rim. Vancouver is well placed to provide these crossroads, where companies from around the world can access new business opportunities. With strong support from trade associations in Asia, Europe, North America, and elsewhere, together with creative input from our Advisory Boards, we are poised to take BPN to a higher level. BPN is a focused event for serious dealmakers and attracts participants from the highest levels of global organizations.

PRODUCED BY: TVG Your Global Life Science Network™ | LifeSciences British Columbia

POWERED BY: biopartnering.com™ contact. connect. collaborate.

HOSTED BY: BIOTECanada | bio alberta

Products, Deployments, and News in Supercomputing

[High Performance Computing] The atmosphere at SC11 was big data-focused and full of new tools.

BY BIO-IT WORLD STAFF

SEATTLE—November's Supercomputing 11 (SC11) conference* in Seattle was full of news and product announcements across the industry. Here's a sampling of what caught *Bio•IT World*'s ear.

Top500 announced their newest list. The Top 10 order didn't change, but the space between systems did. Japan's "K Computer", at the number one spot, completed a build out to make it four times as powerful as China's Tianhe-1A system in the number two position.

The **Cornell Center for Advanced Computing** received an HPC Innovation Excellence Award from the International Data Corporation for enabling hepatitis C virus research on a remote experimental MATLAB computing resource located at Cornell University.

AccelerEyes launched ArrayFire, a freely-available GPU software library supporting CUDA and OpenCL devices. "ArrayFire is our best software yet and anyone considering GPU computing can benefit," said James Malcolm, VP Engineering, in a press release. "It is fast, simple, GPU-vendor-neutral, full of functions, and free for most users."

Jason Stowe (of Cycle Computing) was working the **Amazon Web Services** booth. AWS theorized a bit on the cloud's role in high performance computing in a blog post mid-week, and announced a new type of instance—CC2, or Cluster Compute Eight Extra Large—at the event.

Cycle Computing offered $10,000 worth of free CycleCloud time plus technical support (at SC11 Amazon kicked in their own $2,500) to help a research group address an "un-askable question" as part of its BigScience Challenge. The finalists named at SC11 were Alan Aspuru-Guzik and Johannes Hachmann from the Harvard Clean Energy Project; Jesus Izaguirre from the University of Notre Dame; Soumya Ray from Harvard Medical School; Victor Ruotti from Morgridge Institute for Research; and Martin Steinegger from TU Munich ROSTLAB.

Bright Computing is a company worth watching for their cluster management capabilities in the cloud and new status as an Amazon Web Services Solution Provider. Bright just announced comprehensive support for cloud bursting: with a few mouse clicks, customers

* International Conference for High Performance Computing, Networking, Storage and Analysis, Nov 12-18, 2011, Seattle, Washington

can either create new clusters in the cloud or add cloud-based nodes to existing clusters that Bright manages like part of the local cluster.

Convey Computer announced a partnership with Nimbix (link to PDF) to deliver Convey's unique hybridcore architecture and bioinformatics personalities as a cloud-based HPC solution, and a doubling of its Graph500 performance (link to PDF).

Cray won a University of Illinois' National Center for Supercomputing Applications contract, and will provide a Cray XE6 system and a future upgrade of the recently-announced Cray XK6 supercomputer with GPU computing capability for National Science Foundation's Blue Waters project. Once fully deployed, the system is expected to have a sustained performance of more than one petaflops. Cray also announced the Sonexion 1300 system, an integrated file system, software and storage product.

Along with Cray, **NVIDIA** announced a parallel-programming standard—OpenACC—with the **Portland Group** (PGI) and **CAPS enterprise**. OpenACC allows parallel programmers to provide simple hints—"directives"—to the compiler, identifying which areas of code to accelerate, without requiring programmers to modify or adapt the underlying code itself.

The **DataDirect Networks** booth was busy as the company released new storage offerings tied to its Storage Fusion product line. The company's Storage Fusion Architecture was chosen by Munich-based Leibniz Supercomputing Center to power its SuperMUC HPC system.

Globus Online celebrated its first birthday at SC11, and is doing effective large file transfer. The United Kingdom's National Grid Service has adopted Globus Online as the preferred data movement method for its users.

NETLIST introduced its new 32GB Virtual Dual Rank HyperCloud Planar-X RDIMM, enabling up to 768GB in a standard two-processor server, and demonstrated 1333 MT/s (mega transfers per second) on a standard server.

Oak Ridge National Laboratory, Myricom, and **Juniper Networks** demonstrated "wide-area" 100 Gigabit Ethernet (100 GbE) between the ORNL and Juniper Networks booths. Juniper Networks and Myricom are providing ORNL the infrastructure required to connect to DOE's Advanced Networking Initiative (ANI) 100G testbed, which links several national laboratories.

ScaleMP pushed virtualization with the newest version of its server virtualization for aggregation software platform, vSMP Foundation 4.0, and showcased the product line with partners **HP, IBM, AMD,** and **AdvancedHPC**.

SGI hosted an impressive booth and introduced their Next-Generation ICE X Scale-Out Bladed HPC Cluster. The SGI-Cloudera partnership is also worth watching. SGI will now distribute Cloudera software pre-installed on SGI Hadoop Clusters.

Platform Computing made a big splash, as they always do. Since being acquired by IBM in mid-October, Platform's offerings could get even more interesting.

Bull had a solid presence and launched a new generation of ultra-dense petascale supercomputers, though it's not yet clear how much that will affect life sciences.

Elsevier Acquires Ariadne Genomics

Dutch publishing giant, Elsevier, has acquired Ariadne Genomics, a software provider of pathway analysis tools and semantic technologies for life science researchers, based in Rockville, MD. "Ariadne Genomics' pathway analysis tools and semantic technologies integrate research findings from across multiple content sources providing a deeper understanding of biological pathways and disease progression," commented Alexander van Boetzelaer, managing director of Elsevier Corporate Markets. "Ariadne's products improve research productivity and outcomes for life science researchers by delivering new insights for potential interventions, therapies and cures." The Ariadne acquisition delivers "an information offering in the biology domain and a passionate and dedicated team of life science professionals," van Boetzelaer continued. "Ariadne's team and offerings are a powerful complement to our chemistry, pre-clinical and clinical workflow solutions."

IBM, Pharma Donate Chemical Database

IBM, Bristol-Myers Squibb, DuPont, and Pfizer are providing the National Institutes of Health with a database of more than 2.4 million chemical compounds extracted from about 4.7 million patents and 11 million biomedical journal abstracts from 1976 to 2000. The chemical data should help researchers more easily visualize important relationships among chemical compounds to aid in drug discovery and support advanced cancer research. The data was extracted from patents and journal abstracts using the IBM business analytics and optimization strategic IP insight platform (SIIP, www.ibm.com/gbs/bao/siip), a combination of data and analytics delivered via the IBM SmartCloud, and developed by IBM Research in collaboration with several major life sciences organizations. The platform uses automated image analysis and enhanced optical recognition of chemical images and symbols to extract information from patents and literature upon publication.

NHGRI Pledges $416m to Sequencing, Software

The National Human Genome Research Institute (NHGRI) has announced the latest iteration of its flagship genome sequencing funding program—worth $416 million over four years—that features new initiatives in the search for the underlying causes of rare inherited diseases and accelerates the use of genome sequence information in the clinical arena. "Our goal is to empower all types of researchers and health care professionals to use genomic information in their research and eventually in patient care," said NHGRI


Eric Green

director Eric Green in a press briefing. The bulk of the funding—$86 million in the first year of the plan—will be divided among the three principal genome centers in the United States—the Broad Institute ($35.9 million), Washington University, St Louis ($28.4 million) and Baylor College of Medicine ($21.3 million), under the direction of Eric Lander, Richard Wilson and Richard Gibbs, respectively.

Thomson Reuters Receives Parkinson's Grant

Thomson Reuters has received a grant from The Michael J. Fox Foundation for Parkinson's Research (MJFF) to create and publish the world's most comprehensive source of biological maps for Parkinson's disease (PD). The project aims to identify possible causes of Parkinson's disease by mapping biological mutations of the Leucine-rich repeat kinase 2 (LRRK2) protein, the most common genetic contributor to the disease discovered to date. The maps trace the disease's biological pathways to pinpoint relevant biomarkers, which are biological molecules that signal an abnormal process or disease. They also support the drug discovery process for treating PD.

ChemAxon Moves to Budapest

ChemAxon is moving its European headquarters to expanded offices in Budapest, Hungary. The move will accommodate significant growth over the past few years. "This move marks a key stage in ChemAxon's evolution," commented Alex Drijver, CEO, ChemAxon. "Our new headquarters will enable us to serve our customers better and continue our growth across all aspects of the business."

MMRC Speeding Up Cancer Trials

The Multiple Myeloma Research Consortium (MMRC) has been able to accelerate the clinical trials process by making improvements to trial start-up and enrollment as well as patient accrual for oncology clinical trials run through the Consortium. MMRC presented the data in an oral presentation at the American Society of Hematology (ASH) Annual Meeting. Oncology trials are notoriously difficult to set up, enroll, and complete, often taking as long as two years from design to activation. Between May 2006 and July 2011, 25 multiple myeloma trials were conducted with MMRC project management resources. Trials conducted from June 2006 to September 2008 made up a set of baseline trials, with the subsequent trials serving as test cases for the MMRC's model to accelerate drug discovery. They found that the more recent trials opened 28% faster than baseline trials and enrolled 10% more patients.

The Unstable Equilibrium of the Bioinformatics Org Chart

By Aaron Kitzmiller

In most organizations, the human resources of bioinformatics are a regular source of tension. Unless you're particularly lucky, you can be plagued by politics, illogical decision making, disappointment, and low productivity. While you can have these problems in a properly-balanced organization, there are certain org charts in which they are endemic.

Before discussing the bad organizations, let's briefly enumerate the good ones. While there may have been much groping for solutions early on, modern genome centers and large core facilities consider a dedicated software staff as a given. This full-time, in-house group works exclusively on the software and IT requirements of generating and translating large amounts of sequence data from the instrument into a usable result. Organizational questions about reporting structure, strategic direction, and customer relations are mostly non-existent due to consensus about the nature and scope of the task. The group may be part of a larger informatics organization that provides external services and software, but the team is dedicated.

A genome center can afford to have key technical teams of optimal size (between 5 and 9) that are given relatively focused tasks. These groups are not plagued with the single point of failure and turnover risks of smaller teams and can afford some intra-team specialization (e.g. UI-oriented developer).

At the other end of the scale, a single, skilled individual—more often 'home grown' than recruited— works directly for a lab, with either no other management, or a very weak link to a line manager. This individual can directly solve most data analysis problems through a combination of small databases, workgroup tools, analysis web sites, and downloadable software. Communication difficulties are drastically reduced because there is no development /analysis team and understanding lab personnel and processes quickly becomes internalized. Problems can arise down the road, however, since the individual rarely leaves enough documentation for his or her replacement. While these individuals will need to draw on external resources to accomplish certain tasks from time to time, with the correct mix of friends, consultants, and software vendors, the lab can usually solve a wide range of problems.

The reason these two organizational struc-

tures are successful is that many of the variables associated with bioinformatics service are set to a constant. The intra-team communication and informatician-customer communication issues are set to "1" for the dedicated individual. While the genome center support staff must communicate well internally, their customers and tasks are fairly constant. Far more common are the tense organizations where customer, task, and team are allowed to float.

> Unless you're particularly lucky, you can be plagued by politics, illogical decision making, disappointment, and low productivity.

Informatics Issues

Lab scientist with added bioinformatics responsibilities: In some smaller laboratories, or those venturing into higher throughput sequencing or arrays, a lab scientist with a predilection for technology can begin to take on the automation of data processing and analysis that is overrunning spreadsheets. This quickly becomes a demanding task that competes with lab responsibilities. This was far more prevalent 10-15 years ago when the idea that lab would need software/IT support was novel. The individuals that I've known in this category usually become unhappy because of the variability and uncertainty in day-to-day tasks. This position either decays into a dedicated informatics resource, reverts back to pure lab work when another solution is found; or the individual leaves for a pure lab or pure informatics job.

The Bioinformatics Core: In this setup, some mix of informatics analysts and developers are put together under a common organization that then distributes their time to affiliated laboratories. This makes sense in companies where line management tends to be skill-oriented and is a natural analog to the wet lab core facilities that are commonly established in companies and universities.

But this never works the way people expect it to. Informatics does not fit the core facility mold. One could imagine a disciplined core with a menu of options and strict time frames for them (R scripts for gene expression analysis: 1 month maximum; SOLiD RNA-Seq analysis: 2 month maximum). Coupled with a charge-back system, labs would make careful, focused decisions on narrowly scoped, time-limited tasks. The tasks, at least, would be set to a constant. (Un)fortunately, software development and analysis is so broad that it can be applied to a number of tasks in any given lab. Robotics, LIMS, algorithm development, specialized analysis, statistics, visualization are just of few of the varied dimensions in the bioinformatics universe. A good sized group of 5-10 can cover any of these and more. And charge-back systems are rare, so core staff members are usually "free" resources for the affiliated lab.

The fungible nature of software and analysis and the "free"-ness of the resources com-

bine to drive decisions toward politics, informatics staff interest, and "relevance". There is nothing inherently wrong with an economy driven by favors, but it does not scale. The lab with the largest mindshare and political weight usually gets the lion's share of attention, leading to competition to be associated with the more important projects. Other requests get served in such a way as not to tie down resources for long periods. You can get the core to build a web application to run a custom script, but you can't get the core to do something as general as "improve efficiency by automating critical tasks." As priorities are vague, staff interest will start to drive decisions as well. Vaguely optimistic assignments like "prototype cloud-enabled algorithms," will always win out over critical but unglamorous tasks like "migrate all legacy data into the new LIMS."

A centrally-managed bioinformatics organization can function under different models than the "Core" described above, but organizational fear often prevents adoption. A core with a small menu of well defined services, charge-backs for disk and CPU usage, and a team of scientific consultants can be highly effective. Affiliated labs know well ahead of time what they can expect and what they will have to do for themselves. They also know that the consultant's time is not free. However, as most managers can see, this is a prime outsourcing target. No matter how cost effective and customer friendly the group, some executive looking to have impact will eventually replace this organization with an external company.

The loosely coupled local expert model where the "skilled, dedicated individual" described above is part of a larger bioinformatics organization can work. However it assumes that the bioinformatics management is pretty hands off, doing little more than communicating best practices among the group helping to screen new applicants. Without a capable leader, though, the temptation to absorb the dedicated resources into the lab is very strong. Again, we decay into the "skilled, dedicated individual" system.

Getting the organizational structure around that brain right can be the difference between years of productivity and constant politics, reorganization and lost value.

Informatics support is an increasingly necessary part of laboratories of all sizes. The pace of change has so far prevented the establishment of shrink-wrapped solutions and necessitated the application of the immensely powerful human brain. Getting the organizational structure around that brain right can be the difference between years of productivity and constant politics, reorganization and lost value.

Aaron Kitzmiller is a senior scientific consultant with the BioTeam. He can be reached at aaron@bioteam.net.

The Pharmaceutical Safety Data Problem

By Ernie Bush

I doubt most people fully appreciate the size and scope of the pharmaceutical safety informatics space. I consider myself an avowed information systems advocate with decades of pharma R&D experience, but I did not begin to get a handle on this issue until the last few years.

What we in the industry call "safety data" covers everything from discovery-oriented in vitro or cell based studies (e.g. cytochrome P450 drug-drug interaction results) to extensive and regulated GLP toxicology study data, voluminous clinical study records, and finally to the complex and extensive post-marketing/pharmacovigilance systems. It leads one to wonder: does anyone have informatics systems that allow safety investigators across the pharma enterprise to effectively mine this ocean of information? In my experience the answer is clearly no, or at best only on a very limited scale. This may be a major missed opportunity.

As with many aspects of pharma R&D issues there is both a corporate (or human) aspect and a technology (or systems) aspect. Of these, I have always found the corporate to be the most difficult to address.

Nearly all major pharma companies have three separate divisions that are responsible for some aspect of safety data generation and maintenance: Research (discovery and early development), Development (clinical trials and regulatory submissions) and Marketing (post launch surveillance and pharmacovigilance). These usually report up to the CEO through separate VPs and have different goals, milestones, and bonus/reward structures. No wonder their informatics systems tend to be disparate and lacking good mechanisms to share knowledge or be easily mined across the enterprise.

The solution is probably similar to the one implemented for financial-based issues within the enterprise: have one individual that is accountable for all financial aspects of the business, the Chief Financial Officer. A Chief Safety Officer with the mandate and authority to drive an integrated safety organization would undoubtedly also help drive a more thorough and robust safety informatics strategy.

But the problem is much deeper than just lack of alignment. Perhaps the biggest issue is that neither toxicologists nor clinicians want to open up their data to those "unblessed" members of the company whom they believe

unqualified to interpret the results. There is a fear that individuals without the proper training and background will somehow abuse the information and the professionals would like to limit access to only those they believe will treat the data appropriately. That sense of proprietary ownership between different groups within the same company (sometimes within the same building) is very difficult to address, but will have to change if we are to truly leverage our safety investment to the maximum.

If we mine the entire safety space across the enterprise, would it make a difference?

The Technology Is Here

The problems associated with disparate safety information systems are, in fact, smaller now than five years ago and much, much smaller than ten years ago. And, at least from a technology perspective, it should only get better. The new systems for clinical safety data management and pharmacovigilance are merging (or at least becoming more aware of each other) and the tools for accessing data from multiple legacy systems have blossomed.

Ultimately, however, the big question remains: If we truly could mine the entire safety space across the enterprise, would it make a difference? Traditionally the answer was believed to be no, or at least not very much. There has always been a sense that the clinical study data "trumps" the other data and therefore other data—especially preclinical—does not add a lot to late stage or post-marketing safety assessment. And from a traditional "observational" based safety assessment point of view, that seems logical.

But increasingly, safety assessment is moving to a biochemical- and mechanism-based approach. And here, all three legs of the safety informatics space really carry an equal amount of the insights. After all, it is in the preclinical area where we develop the best understanding of the mechanisms of adverse effects. For example, currently most companies screen against 800 to 1,000 off-target receptors in order to get a feeling for what kinds of adverse events could be seen in the clinic. This is very valuable data, yet how many post-marketing surveillance organizations can explore whether there was an activity at an off target receptor that might help manage an adverse event in a new population? Actually, a few do and the list will grow more, which proves the point.

An enterprise-wide safety assessment informatics solution is the optimal tool for fully leveraging the investment in generating and collecting this data—especially to achieve a mechanism-based understanding of what is behind the safety issues. This in turn greatly improves the ability to manage or mitigate these adverse effects.

Ernie Bush is VP and Scientific Director of Cambridge Healthtech Associates. Email. ebush@chacorporate.com

10th ANNIVERSARY 2002-2012

Cambridge Healthtech Institute's Eleventh Annual

Bio·IT World CONFERENCE & EXPO '12

April 24 – 26, 2012 • World Trade Center • Boston, MA

Enabling Technology. Leveraging Data. Transforming Medicine.

CONCURRENT TRACKS:

1. IT Infrastructure – Hardware
2. IT Infrastructure – Software
3. Cloud Computing
4. Bioinformatics
5. Next-Generation Sequencing Informatics
6. Systems and Multiscale Biology
7. eClinical Solutions
8. eHealth and HIT Solutions for Personalized Medicine
9. Drug Discovery Informatics
10. Molecular Diagnostics Informatics NEW!
11. Open Source Solutions NEW!
12. Cancer Informatics NEW!

EVENT FEATURES:

- Access All 12 Tracks for One Price
- Network with 2,000+ Global Attendees
- Hear 125+ Technology and Scientific Presentations
- Connect with Attendees Using CHI's Intro-Net
- Choose from 16 Pre-Conference Workshops
- See the Winners of the following 2012 Awards: Benjamin Franklin, Best of Show, and Best Practices
- View Novel Technologies and Solutions in the Expansive Exhibit Hall
- And Much More!

KEYNOTE PRESENTATIONS BY:

Eric D. Perakslis, Ph.D.,
CIO and Chief Scientist of Informatics, U.S. Food and Drug Administration

Martin Leach, Ph.D.,
CIO, Broad Institute of MIT and Harvard

Jill P. Mesirov, Ph.D.,
Associate Director and Chief Informatics Officer; Director, Computational Biology and Bioinformatics, Broad Institute of MIT and Harvard

KEYNOTE PANEL:

A special plenary session featuring trends and challenges in cancer research:

Julian Adams, Ph.D., Vice President, Business and Corporate Development, Infinity Pharmaceuticals, Inc.

Jose Baselga, M.D., Ph.D., Chief and Bruce A. Chabner Chair, Division of Hematology/Oncology, Massachusetts General Hospital; Associate Director, Massachusetts General Hospital Cancer Center; Professor of Medicine, Harvard Medical School

John Quackenbush, Ph.D., Professor, Biostatistics and Computational Biology, Cancer Biology Center for Cancer Computational Biology, Dana-Farber Cancer Institute

Platinum Sponsors:

华大基因 BGI

BLUE ARC PART OF HITACHI DATA SYSTEMS

BT

CYCLE COMPUTING

hp

ISILON

panasas

Official Publication:
Bio·IT World

Register by January 27 and Save up to $500!

Please mention keycode U02 when registering!

Bio-ITWorldExpo.com

Organized by:
Cambridge Healthtech Institute

Featured Speakers

Pankaj Agarwal, Ph.D., Dir Computational Biology & Bioinformatics, GlaxoSmithKline

Dimitris Agrafiotis, Vice President, Informatics, Johnson & Johnson PRD

Garen Avetissyan, Senior Manager, Clinical Business Systems, Biogen Idec

Toby Bloom, Ph.D., Director, Informatics, Genome Sequencing, Broad Institute

Professor Sir John Burn, M.D., Fellow of the Academy of Medical Sciences, Professor of Clinical Genetics, Newcastle University

Michael Cantor, M.D., Senior Director, Biomedical Informatics Services, Pfizer; Assistant Professor of Medicine, NYU Medical Center

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

David Dooling, Ph.D., Assistant Director, The Genome Center, Washington University, St. Louis

Ramesh Durvasula, Ph.D., Director, Molecular Sciences & Candidate Optimization Informatics, Bristol-Myers Squibb

Christine Gibson, Associate Director, Clinical Business Systems, Biogen Idec

Stephen Granite, Director of Database and Software Development , Center for Cardiovascular Bioinformatics and Modeling , Institute for Computational Medicine , The Johns Hopkins University

Andrew Kasarskis, Ph.D., Vice Chairman, Department of Genetics and Genomic Sciences; Co-Director, Institute for Genomics and Multiscale Biology, Mount Sinai School of Medicine

Isaac Samuel Kohane, M.D., Ph.D., Director, Children's Hospital Informatics Program, Henderson Professor of Pediatrics and Health Sciences and Technology, Harvard Medical School

Kevin Krenitsky, M.D., Chief Operating Officer, Foundation Medicine

Anthony Leotta, Bioinformatics Manager, Cold Spring Harbor Laboratory

Roger (Rangjiao) Liu, Ph.D., Research Manager, Bioinformatics, Life Science Development, Corning

Subha Madhavan, Ph.D., Director, Clinical Research Informatics, Lombardi Comprehensive Cancer Center; Director, Biomedical Informatics, Georgetown-Howard Universities CTSA, Georgetown University Medical Center

Keith Perry, Associate Vice President & Deputy CIO, University of Texas MD Anderson Cancer Center

John Quackenbush, Ph.D., Professor, Biostatistics and Computational Biology, Cancer Biology Center for Cancer Computational Biology, Dana-Farber Cancer Institute

Vishal Rosha, Senior Scientist, Bioprocess Research & Development, Novartis Pharma Ag

Anthony Rowe, Ph.D., Principle Scientist - External Innovation, R&D IT, Janssen

Sandeep Singhal, Bioinformatician, Breast Cancer Translational Research Laboratory J.C. Heuson (BCTL), Jules Bordet Institute

Vas Vasiliadis, Director, Products, Computation Institute, University of Chicago

Xiaoming Wang, Ph.D. Fellow, Computation Institute, University of Chicago and Argonne National Laboratory

Zheng Yang, Ph.D., Associate Director, IS Business Partnering Research, Boehringer Ingelheim Pharmaceuticals, Inc.

Alexander Wait Zaranek, Ph.D., Director, Informatics, Harvard Personal Genome Project, Genetics, Harvard Medical School

For a complete list of speakers and topics, and to register, visit www.bio-itworldexpo.com

Featured Topics

Collaborative Drug Discovery

Drug Repositioning

Health Informatics Networks in Clinical Decision Support

HIT & Personalized Medicine & Health: Real World Use of EHR Data

Using Data Visualization to Improve the Efficiency and Quality of Clinical Trial Monitoring

Overcoming Challenges in Integrating Systems (EDC, IVRS, CTMS) and Clinical Operations Data

Moving Beyond the Mean: The Role of Variation in Determining Phenotype

Tying Individual Molecules to Phenotype

Technologies and Infrastructure for Acquiring, Storing, and Sharing Data

HPC on the Cloud – Implementation Issues

Open Source Platforms and Collaborative Technologies

Data Modeling and Computational Integrative Tools

Translating Genomic Data and Research into Clinical Practice

IT Infrastructure on the Cloud

Hybrid Strategies for Biomedical Research Data Management

Clinical Application of Molecular Diagnostics

Multiplex Assay Accuracy and Test Result Interpretation

Pre-Conference Workshops

Morning Workshops

Beyond the Cloud

Leveraging Saas for Next-Gen Sequencing

Utilization of EHRs/EMRs for Protocol Design

Building and Using an Ontological Framework

Clinical Genomics

Collaborative Innovation and Open Source Deals

Marketing and Sales: Science Training 101

Best of Breed Informatics

Afternoon Workshops

Data Visualization in Biology

Microscopy Imaging Analysis

Cloud Computing

Enhancing R&D Effectiveness through Global ELN Deployment

Finding eHealth Savings

Seeking Precise Knowledge

Social Media Analytics

Scientists: Business Training 101

Integrating Infrastructure for Clinical Trials

Medidata on the cloud, 'omics, and personalized medicine.

Medidata co-founder Glen de Vries

Glen de Vries thought he would be teaching biology or chemistry in college, but somewhere en route to a satisfying career in academic research, he got distracted. He became exposed to clinical research while working on a molecular biology project at the Columbia Presbyterian Medical Center. Together with a friend and medical resident in the same lab, he ended up quitting to pursue Medidata Solutions Worldwide full-time. De Vries is currently president of Medidata Solutions, a provider of clinical technology solutions that believes in "optimizing clinical trials from concept to conclusion." He was the original architect of the Medidata Rave electronic data capture (EDC) and clinical data management system (CDMS).

Bio•IT World chief editor **Kevin Davies** spoke to de Vries about the progress of Medidata and the state of e-clinical technology in general.

***Bio•IT World:* Glen, what was the original goal of Medidata?**

de Vries: Medidata was founded because we'd been talking over ideas about taking data management during a clinical trial to the Web. We found someone who thought it would be a great idea for a clinical trial he was running. We wound up taking the bed out of my one-bedroom apartment and converting it into a studio apartment with an office, and that's how the company was born! The sponsor of the trial, when they heard that our systems were being run out of a closet, said that makes no sense. We've got to be serious about creating an infrastructure that people would trust. That was in 1997.

By 1999, we had all our ducks in a row and really opened the doors on Medidata. At that point, we met Tarek Sherif, the other co-founder [and chairman/CEO]. Culturally, it's still useful to think [the way we thought back then]—a combination of thinking about basic science, practical aspects of providing health care, being a clinical trial site and a sponsor, and a healthy dose of business realities. We're not just a technology vendor but one working in the world of services that get outsourced.

Would you call yourselves a CRO (contract research organization)?

We are not a CRO. In its most extracted pure form, we're trying to provide a platform—the technology infrastructure—that pharma, biotech, and medical devices will need in the future. I'd stress that we're trying to build for the future state.

Over the last 15 years, most people have approached clinical trial technology infrastructure by taking individual paper processes and automating them. Sometimes you can make it more efficient, but nobody has really tried to take an approach similar to the ERP (Enterprise Resource Planning) industry. If you look outside life sciences, companies would have different departments to manage individual processes (manufacturing, sales, customer support etc.), then companies like SAP developed platforms that fall under ERP—bringing all the functions in a company together in one central software base (e.g. ordering a custom computer over the Internet). There hasn't been an ERP-like platform in clinical development.

We see that as an opportunity—instead of replacing individual systems, many of which are 30 years old, we're trying to think about how to connect all the different people, from the executive in charge of an entire drug program or research portfolio to the patients who volunteer to be the subjects in a clinical trial. Can we create much more modern, efficient business processes in a clinical trial that demonstrates safety and efficacy, present it to a regulator, and get permission to bring it to the marketplace?

How does that translate into your core mission now?

We're trying to build that platform to accommodate as many of our customers' needs today. If you look at Rave, which is one of our core solutions, there's an interesting illustration of that idea. We never bought into the idea that an EDC system that handles the data in a CRF [case report form] should be separate from a clinical data management system that integrates CRF data with lab data and with randomization data. We thought that should be one system, and that's what we built. We've gotten people in the industry to change their expectations around that. It used to be that people said, "Those things used to be two things. I can't manage all that data." Well, it is 2011. I think you can manage that way!

Our customers who are doing this are creating much more efficient processes, getting visibility into their data earlier. And we hope to deliver for them a real competitive advantage.

When we think of a technology stack or a business process... we see ourselves as the bottom tier of a stack of layers focused on running the most efficient clinical trial today and in the future. We try to support our sponsors and CRO partners by getting their people on top of that platform.

How much have you been able to integrate into your platform and where are some of the remaining holes or opportunities?

We don't see any finish line where we'll be done with our platform. The idea is to always think of what people will need in the future and to invest in things that will be useful for those clinical trial processes. If you look at companies in a similar position to us, you'd see something like five percent revenue being put back into R&D. We put more than 15 percent revenue back »

into R&D. We see so much change coming in managing the clinical development process that we want to make sure we stay on top of those future requirements.

How far have we come? With Rave, we started to deliver EDC and CDMS in a new, hopefully better, way. To categorize some of the system types, we've added randomization and drug logistics [IVRS]. We've recently added capabilities in CTMS (clinical trial management system). We think we're defining a category that didn't exist in terms of structured protocol design—leveraging a database system to ensure you're designing the best protocol possible, and using that structured design to inform other activities, like budgeting the clinical trial and managing the way you deal with site grants.

We're constantly looking for opportunities to integrate technologies we don't have. We talk about 'new, different, and better'—it's our job to show it's better.

Do you typically build new capabilities internally or acquire new technologies as you need them?

We build a lot, and building technologies is one of our core competencies. We think software development is an art form and we're proud of the way we do it. But we will, and just did, make a tuck-in acquisition, if we find something compatible philosophically with the technologies we have to date.

We just acquired a company called Clinical Force, which has capabilities in the CTMS category. There's a philosophy we have—that we applied to Clinical Force and Fast Track Systems, another company we acquired—that we're not interested in just assembling a suite of software that meets industry requirements.

> "Instead of replacing individual systems,... we're trying to think about how to connect all the different people, from the executive... to the patient.

It must be something that provides value in the platform context. Each individual part needs to add value if used alone, because it's a great EDC system or investigator payment administration system, but, when used in conjunction with the other parts of our platform, it has to be part of a seamless flow of data between the pieces of the platform. We're big believers in open interfaces, so when we build or integrate an acquired technology, the way we connect them is with the same set of tools we provide to partners and effectively the general public. We have a website called Developer Central, where anyone can get access to all our documents and APIs. Those are the same tools that our internal developers are using... That openness is key. How can we make any new system part of our platform in a very holistic way?

Any other significant new acquisitions we should be aware of?

Clinical Force has put us into the CTMS space. The way CTMS has been implemented by many sponsors, it's often provided on top of a heavyweight, legacy on-premise system, often highly customized, very costly to upgrade. A lot of people are talking about a better business process to attack trial management activities,

requiring a change to the CTMS. But the systems people think it would be too expensive or time-consuming to upgrade their CTMS. My thought is, isn't the technology supposed to be there to make it easier to implement a better business process? We think, "Great, here's an opportunity."

We're firm believers in software-as-a-service—not just hosting the application for customers, but making it so they don't have to worry about implementing the application. We try to make everything as self-service as possible. Our approach to CTMS is to provide it in a similar fashion—consume the modules you need, when you need them, easily over the web. If you're using them with other systems, use a set of open interfaces.

What is your stance on cloud computing? Are clinical data incompatible with the cloud?

If you're looking at the world of clinical development—I wouldn't just place Medidata into this category—we are the cloud for clinical data. In the very first days of Medidata Solutions, when we first started doing EDC, the vast majority of sponsors and CROs had no infrastructure to run an over-the-firewall e-commerce application. Their networks were buttoned up because they were used to running in a very conservative pharma-IT environment.

From the first days, we were in the business of hosting EDC, then other systems, for our customers. They just access them over the web. The [clinical] sites' doctors and nurses access the software over the web, then via the web they download the data. That is effectively the cloud business model.

The term cloud is thrown around a lot in various contexts in marketing and other agendas. We think today the cloud is business as usual for us, as it has been for 12 years. We think the way the entire world of IT is going is based on the same concept. The end user shouldn't have to worry about the infrastructure being used to deliver the functionality you need; you should just be able to consume that functionality.

Here's an analogy: Income tax software started with a paper process. You'd have to sit down with your receipts and transcribe onto paper, just like a CRF. So what happened? People sold you software to transcribe your data into a computer... Now, you don't even have to buy a CD anymore—you just go to the web and purchase the functionality you need and enter your data. That's become a cloud service. Putting your information into these sites means the software you're interacting with can give you real-time feedback on how well you're doing... It's leveraging a much larger body of data.

This is not so different from how we think clinical data should be managed. We started with paper, people started building heavy client software, it moved to the web... We've been showing demonstrations on how to load data from electronic health record systems into the Medidata platform, effectively bypassing the EDC data entry process entirely. What better way to collect clinical data than to have it happen that way?!

It's fair to say that big pharma's productivity has been pretty mediocre for the past several years. What is your view on the potential impact and application of new technology, including 'omics technology, on improving

clinical research through things such as biomarker discovery and patient stratification?

Let's start with the availability of real integrated platforms such as ours—this will help people get much more effective at the general people-based processes of designing and managing clinical trials. There is just a spectacular opportunity for making this not just more efficient but more effective.

When ERP came around, the next thing that happened was everybody started trying to design new business processes based on the new capabilities provided by these new technology platforms. Then they realized they needed to effectively measure how well those new processes were working. A whole industry of business intelligence was born, based on that end user requirement. People wanted to have the right dashboards and predictive capabilities to manage their businesses much more aggressively.

Right now I think we're living through the same transformation in clinical trials. If you look at the huge amounts of incremental data that we think will help life sciences get new drugs approved, or new combinations, we can't just throw resources at the problem and expect it to work. We're talking about an exponentially increasing amount of data we need to wrap our minds around and associate with labeling claims we want regulatory agencies to approve. There is no way to add that capacity at our current level of efficiency.

We need to use more sophisticated transactional capabilities. We need to layer on the right analytics—we feel passionately about benchmarking as the key to analytics—so companies can look at their own performance, compare their performance to peers and get more effective. Then we'll have the ability to put more and more data through the clinical research pipeline and answer more sophisticated questions.

But what about personalizing medicine through new technologies and strategies?

From the perspective of a patient getting better medical care, personalized medicine is inherently a great idea. As the industry takes the words personalized and medicine though, the two ideas don't scale very well in their current state. Now we're trying to put them together, we know we need to do that, but it won't be possible until we really wrap our minds around how to leverage technology.

If we do that, there's an exciting, healthy future for personalized medicine and leveraging new genomic data. If you look back at the Human Genome Project, there was a somewhat naïve view in the general public that all of a sudden, we could understand and cure a lot more diseases. The reality is a lot more complicated. Just having access to a ton more data doesn't make the data more useful. The relationship between genotypic data and phenotypic data and the vast range of combinations for any given patient mean this is a really different class of problem. I think we'll be able to equip people to tackle those data but we need to start with basic blocking and tackling.

Is the ROI Medidata provides principally saving sponsors money, increasing efficiency, expediting trials, or failing ineffective drugs sooner?

We see ROI coming from two main angles. First, if you're spending money on Medidata technology, you should be able to save money in the

operational component of your clinical trial. For every dollar spent on technology in a clinical trial, there's more than 25 spent on the operations and supplies required (manufacturing test drug, distribution, monitoring, compensation etc.). Our premise is, if we can help redesign your processes so steps can be automated, or change other business processes, there's definitely the financial ROI.

But if you think about changing business processes, eliminating or automating steps, another key ROI is the speed by which you'll be able to access and leverage the information you've collected. You'll be able to do an interim analysis that much faster, or identify safety or efficacy clinical data earlier to figure out where you might have a problem or whether to change course. From an operational perspective, how quickly can I identify sites that are giving me high-quality data and other sites giving me low-quality data that I might want to spend more time on?

> "The reality is... just having access to a ton more data doesn't make the data more useful.

The third piece of ROI is harder to measure in a tangible time or dollar sense, but it's being ready for what is going to be the future state of clinical trials. If you look to the forefront of the commercial and academic worlds, they are not just thought leaders but they are designing and running studies that are incredibly progressive in terms of finding combinations of therapies that are most effective, looking for ways to expose as few patients as necessary to harmful doses, etc. These are studies that are difficult or impossible to execute if you don't have the right kind of integrated infrastructure. I cannot scale up commercially doing hybrid Phase II/III studies if I've taken a traditional view of how clinical trial systems work. So we think a key ROI with Medidata is—by putting your operation on top of our platform, we're there to help make sure you're ready for the future state. And we look to our customers to help us make sure we know what that future state is going to be.

Any final thoughts that we haven't touched on?

Yes. There's an interesting dynamic in what's happening in higher echelons in clinical trial management that sit above Medidata right now. We think about the technology infrastructure. If you look at what's happening in the sponsor/CRO world, there's interest in more and more outsourcing, so CROs are getting more aggressive about distinguishing their offerings. Can a CRO provide value beyond just executing on traditional requirements of data management or clinical monitoring? I think that's a really exciting thing in our industry, because when you shift responsibilities between two parties, it creates an environment where people are looking for new ways to provide that service. That's dynamic and we think that's great for our business.

Hopefully the result is sponsors getting their trials done more efficiently, running their programs faster. CROs are building healthy businesses and looking to distinguish themselves. When you get to the bottom of this, there are companies like us benefiting from people demanding more and better functionality.

A QuantuMDx Leap for Handheld DNA Sequencing

From pharmacegenomic testing to handheld sequencing within 5-7 years?

BY KEVIN DAVIES

MONTREAL—Speaking for the first time in his life as a commercial consultant rather than a public servant, Sir John Burn, a highly respected clinical geneticist in the United Kingdom, provided the first glimpse at a nanowire technology for rapid DNA genotyping that could eventually mature into the world's first handheld DNA sequencer. Burn previewed a potentially disruptive genome diagnostic technology in a presentation on the closing day of the International Congress of Human Genetics in Montreal*.

Sir John Burn unveiled nanowire sequencing technology from QuantuMDx.

One day a week, Burn, professor of clinical genetics at Newcastle University, serves as medical director for QuantuMDx (QMDx), a British start-up co-founded by molecular biologist Jonathan O'Halloran and health care executive Elaine Warburton.

The potential of nanowires was demonstrated in a 2001 *Science* paper authored by Harvard University's Charles Lieber and colleagues. That proof of principle, according to Burn, showed that changes of impedance in silicon nanowires could record the arrival of a biomolecule. Recently, QMDx announced the exclusive license of intellectual property from Lieber's company, Nanosys, of the diagnostic and sequencing applications of nanowires and nanotubes. (The arrangement succeeds a non-exclusive IP deal previously announced in 2009.)

QMDx operations and manufacturing are currently headquartered in Newcastle, U.K.. The company has built a prototype DNA sequencer, which Burn admitted didn't exactly conform to handheld dimensions just yet. (The company is working on a sequencing device called Q-Seq, but Burn did not dwell on that.)

Pass the Q-Poc

The first goal is a handheld point-of-care instrument that is dubbed Q-Poc, which Burn likened to "a big fat iPhone." Possible applications range from companion diagnostics and genetic testing to DNA forensics. The QMDx technology uses

* 12th International Congress of Human Genetics, October 11-15, 2011, Montreal

peptide nucleic acids (PNAs) to tether DNA sequencing templates to a nanowire. As the nucleotides arrive at the template, their negative charge influences the nanowire and changes its impedance. The process would use a reversible-terminating technology that is able to read through repetitive stretches of DNA.

Burn acknowledged that the electrical differences could only be detected for a range of about 100 Angstroms (10 nanometers), equivalent to about 50 bases of DNA in any one stretch. But down the road, an idea was to lay a longer DNA template, say 10 kilobases, and interrogate at multiple points along the molecule.

The manufacturing process relies on an etching technique featuring a silicon nanowire about 120 nm in diameter, and about one million features on the chip, roughly the size of a microscope slide. "It's a highly reproducible system for mass manufacture," says Burn.

To place the DNA templates onto the wire, QMDx has invented several methods to rapidly prepare and PCR DNA. A DNA extractor uses a nanoparticle filter to capture and elute the starting DNA. A microfluidic cartridge performs PCR (polymerase chain reaction) not by heating and cooling the liquid, but by passing it through heat zones created in the microfluidic device.

"There are up to 32 silos of passage across the heat plates," Burn explained. "We can do PCR in 6 minutes through this device—10 minutes from sample arrival to PCR product coming onto the nanowire."

Eventually the sample prep segments will be combined to produce a disposable chip with on-board chemistry that is inserted into the machine. The results can be read locally or transmitted to a remote reader.

Burn listed several early applications that QMDx is targeting. The first is the pharmacogenomic polymorphism that controls sensitivity to the blood thinner warfarin. Today, Burn said, "We give all [patients] the same dose; the ones that nearly bleed to death, we bring them back to hospital and transfuse them and put them on the right dose. It would be nice if we could do the genetic analysis and predict the right dose instead." Using an algorithm developed by his Newcastle colleague Ann Daley, Burn wants to use warfarin testing as "a proof of principle for the use of this as a point-of-care test in clinics and GP surgeries."

Another application is in the field of bowel cancer, a long-time research interest of Burn's. The detection of key microsatellite (MSI) risk markers should be identified at the point-of-care, said Burn. "We'll be moving into the pathology lab and doing genetic testing of MSIs, even as samples come out of the clinic or even surgery." Burn even suggested the possibility of using the technology during live surgeries, as the process should only take some 15-20 minutes.

In short the goal is to carry genomics into the world, Burn said. The 20-minute test should prove valuable for pharmacies and hospitals. The device would be priced in the $300-400 range, with the disposable cartridges about $20. The company plans to begin clinical trials for HIV testing in South Africa in the next 1-2 years.

"I'm convinced—and obviously hope—the nanowire idea is coming to a clinic near you," said Burn. "We're limited more by our imagination than anything else." Burn's QMDx colleague, chief science officer Jonathan O'Halloran, goes even further. "I predict that we'll be able to sequence a genome on a handheld device within 5-7 years," he writes on his LinkedIn page.

Cheminformatics in the Cloud

Accelrys launches next-generation informatics suite and cloud portal.

BY KEVIN DAVIES

Accelrys announced its next-generation informatics suite at its European user group meeting in Athens last year. The suite consists of some updated existing products and some new components, explains Accelrys' senior director for life science marketing, Rob Brown. Those features include a chemical registration system, a chemical data cartridge, and a cloud-based product called HEOS, in partnership with SCYNEXIS, for externalizing research.

"We're building on the joint heritage of Accelrys, MDL and Symyx—more than 25 years," says Brown. "This is the first time there's been a coherent suite to cover all capabilities that scientists need to cover daily workflows," established through a combination of in-house development with acquisitions and partnerships.

"Data gathering steps can be very time-consuming," says Brown. "Scientists should be spending time doing innovation." During a recent case study, Brown says a big pharma partner plotted the time taken by its scientific teams during various phases of drug development. "What alarmed the pharma was that a particular compound was made very early, but a lot of results were locked up in labs doing various tests. The informatics systems not sufficient to get results back to the team," says Brown.

Meanwhile, many other compounds have been screened and tested, which could amount to wasted effort. "Now they have to go back and change direction. Tests can be done quickly, but the informatics systems can't deliver information."

Something New

The Accelrys Cheminformatics Suite is designed to gather data more efficiently, leading to better informed decisions early on while designing compounds, says Brown.

Complementing existing commodities such as Isentris are new features such as a chemical registration system—the first new cheminformatics product developed since the merger between Accelrys and Symyx. "It lets the medicinal chemists register their molecules into the corporate database. It's a modern, Web-based product that should be very seamless for the chemists to use," notes Brown.

Accelrys has also upgraded the companion biological registration system. "By having both registration systems on the same architecture, as we go forward, we can think how to handle those situations when you're registering an entity with attributes of both a biological and a small molecule," Brown says.

Additional updates have been made to the Accelrys chemical cartridge, which allows users to store data in an Oracle database. "Now we can handle patent structures and chemically modified biologics," says Brown.

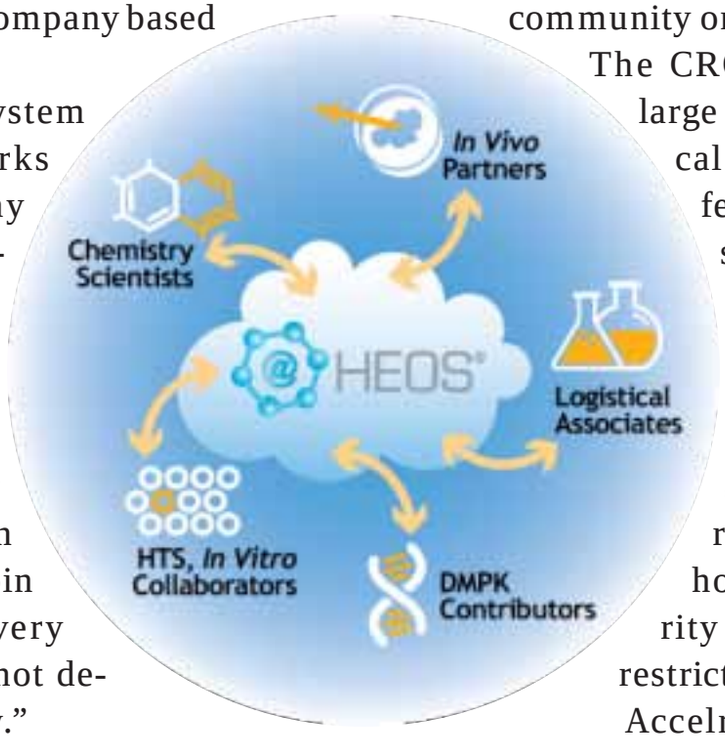
A key question is how to represent a chemical entity in software so it's a faithful reproduc-

tion of the structure. "Many companies offer cartridges to handle regular small molecules... We now have single representation not just for small molecules but reactions, patent structures, and chemically-modified biologics. So it's a much broader representation." And Brown points out, there are tens of thousands of users already.

Holy HEOS

Perhaps the most exciting element of the new suite is the integration of HEOS, a cloud-based platform developed by SCYNEXIS, a drug discovery and development company based in North Carolina.

As the pharma ecosystem evolves, partner networks sometimes include many discovery partners, academic groups for safety, and CROs. "Partners come and go quickly," says Brown. "I might use six partners, but only 2-3 are active at any given time. We might need to spin a partner up or down very quickly. IT in pharma is not designed to that very quickly."



HEOS is a cloud-based platform to manage partners and projects.

The cloud offers an attractive solution in principle, and Brown notes that attitudes to the cloud are changing: "Fears of data security are diminishing. It also fits the operating expense model. The budget comes from business, not IT."

Accelrys has signed a partnership with SCYNEXIS for its HEOS software, which is hosted in a data center managed by the CRO. "They wanted to add value for large pharma," says Brown. "They built this system to handle all types of data." The closest competition, as he sees it, is from Collaborative Drug Discovery.

Users will subscribe to a software-as-a-service model. "We spin up a big sandbox for the operation and partition areas for each of the CROs," says Brown. "The commissioning company decides who gets to see what. Importantly, it's all a real-time upload, making data available to the people who need to see it."

Brown says the value of HEOS has already been demonstrated in the neglected disease community on multi-continent projects. The CRO has passed audits for large pharma, including an ethical hacking audit. (Security features of HEOS include a state-of-the-art data center with physical security features and 48 hours of power autonomy; off-site encrypted data backup and geographically separated disaster recovery; a host of web/software security measures, firewalls, and restrictions on incoming IP.)

Accelrys will license seats for HEOS and/or the registration. "Some start-ups might say HEOS is their internal system, and just license seats to that," says Brown. As for the new suite, users can pick and choose their desired elements. "In most companies, pieces of the [Accelrys] suite will already exist," says Brown. "We can't always replace everything—we have to offer say our chemical registration, but the user continues to use someone else's assay data management tool, for example."

Remedies for SAFER DRUGS

[Drug Safety] New solutions to seek signals, organize data, train interpreters, expedite reporting, and find better biomarkers.

BY DEBORAH BORFITZ

Pharmacovigilance experts have an abundance of signal detection tools to sift through large quantities of data seeking causal relationships between adverse events (AEs) and experimental drugs. They also have an assortment of data mining tools capable of finding statistical associations suggestive of problems regarding approved drugs. All this technology is intended to safeguard clinical trial participants, patients, and the reputation of recall-weary drug developers. But drug safety specialists can't be sure which technology or signal detection method is best.

Drugs today aren't as safe as they could be for several big reasons, including failure to see the big picture, says Steve Jolley, principal of SJ Pharma Consulting. "It took years to figure out Vioxx was killing... 2-3 patients per thousand due to heart attacks and strokes because the people taking it were likely to die of [those

causes] anyway."

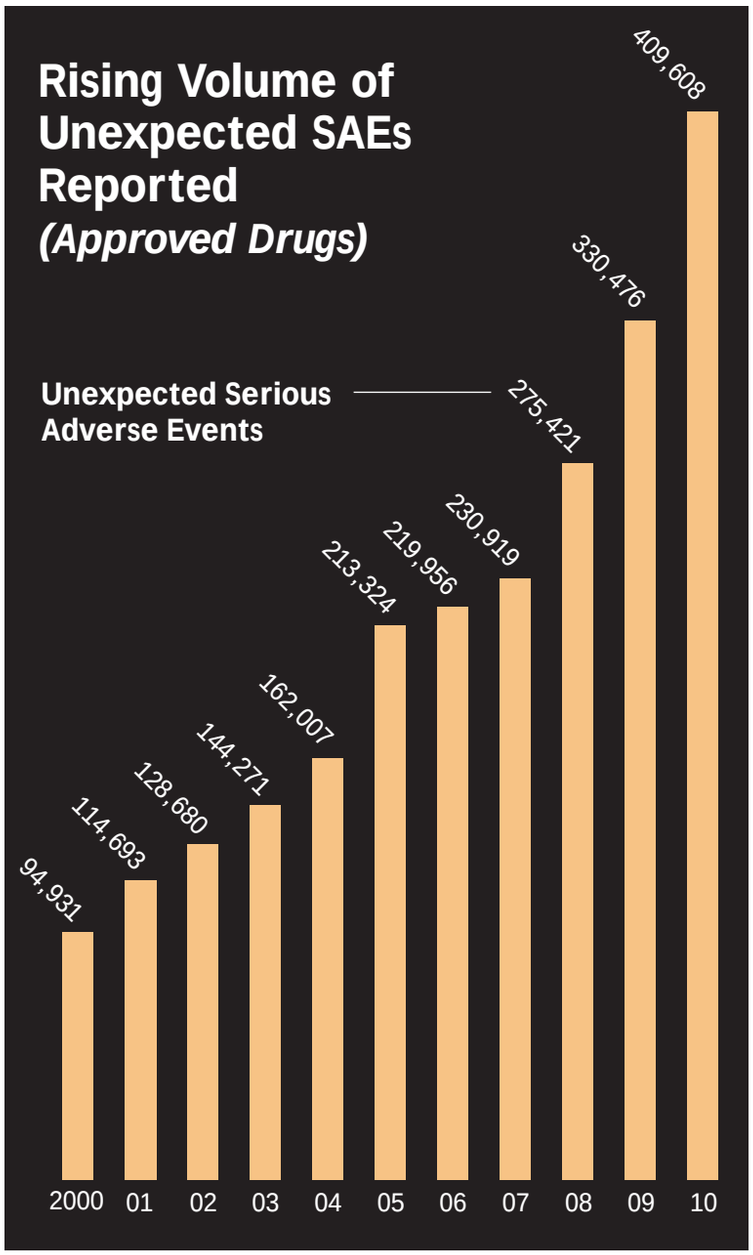
Software tools by themselves are of little consequence without professionals medically interpreting the signals to determine which ones are legitimate product safety concerns, says Sally Van Doren, president and CEO of the drug safety consultancy BioSoteria. But there is a shortage of trained drug safety practitioners evaluating product safety signal detection, she says. "There's a huge training need in this area. Pharmacovigilance, risk management, and signal detection training—including certificate programs offered by organizations like BioSoteria and the Drug Information Association—are a good start but not completely sufficient," because certification confirms only class attendance. No competency testing is presently available or required for those actually entrusted with drug safety surveillance. The proficiency of plumbers and accountants is better defined!

Post-marketing regulatory inspections provide some benchmarks about how well product safety signals are detected, evaluated, and acted upon by pharmaceutical companies, she says. For example, the UK Medicines and Healthcare Products Regulatory Agency's latest metric report of key pharmacovigilance inspection findings notes limited safety signal detection systems in about 13 percent of pharma inspections.

Heightened recognition of the importance of safety signal detection and evaluation has improved therapeutic risk management. Drug makers can ill afford suboptimal patient outcomes and the expense of legal damages as a result of shortcutting safety monitoring. So if preclinical or early-stage clinical studies suggest an experimental drug might be associated with kidney or liver damage, additional biomarker measures get monitored in addition to standard blood and urine tests. The expanding field of pharmacogenetics has also identified genetic differences in patients with AE susceptibility.

Signal Detection

Signal detection of common AEs is best accomplished in a controlled clinical trial setting where the sponsor can collect details of reported AEs and calculate actual incidence. Safety problems are more likely to be detected as data accumulates and clinical development progresses, »



SOURCE: FDA

often with thousands of patients with the targeted disease, says Van Doren. Even in blinded studies, a data safety monitoring board often reviews incoming safety data for imbalances in the drug's safety profile compared to placebo or marketed therapy.

Still, safety is a hard endpoint to evaluate relative to efficacy, says Cynthia Uber, former VP medical services at Eisai. "Safety requires exposure to broader populations than are utilized in clinical trials and can therefore be more costly and time consuming to confirm. It is also subject to a labyrinth of reporting requirements that can become labor intensive," she says. "Safety centers are typically not revenue generating for pharmaceutical companies."

One challenge for safety departments in past years was that those evaluating and reporting serious adverse event (SAE) data often had little access to the full clinical data set during a clinical trial, with separate databases housing different safety information. Today, with electronic data capture (EDC) systems, sponsors may interface the clinical data with a safety database, making "real time" SAEs and patient data available for analysis. Most sponsors also utilize alert systems triggered from clinical databases and sent to the safety monitor when individual patients experience predefined thresholds of certain AEs or abnormal laboratory values, Van Doren says (see "FDA Final Rule").

The Weaknesses of AERS

FDA currently accepts only post-marketing safety reports in its Adverse Events Reporting System (AERS). But the agency anticipates electronic transmission of—and presumably an electronic repository for—IND safety reports on experimental products. In Europe, expedited safety reports of serious unexpected adverse reactions are amassed in the EudraVigilance database for both investigational products and

FDA Final Rule

Effective March 2011, the FDA's Final Rule on IND Safety Reporting required study sponsors of drugs and biologics to look periodically at individual and aggregate safety data from all sources. If any safety finding indicates a significant risk to patients, FDA and participating investigators must be alerted expeditiously.

In past years, clinical trial investigators and institutional review boards (IRBs) were inundated with reports from sponsors of unexpected SAEs, says Sally Van Doren. But the accumulation of such individual patient reports lacked cumulative experience and aggregate safety analyses for investigators and IRBs to fully interpret. If patient-level safety information is going to be of practical value, it should be organized in an accessible electronic data repository by the sponsor.

"Just staring at raw individual patient safety data, without analysis, is no way for the sponsor, investigator, or IRB to protect patients," says Van Doren. Building a real-time central repository for integrated safety data across multiple trials of a drug will require "savvy" individuals to design the database, ensure data collection consistency and determine data outputs, she says.

post-authorization (marketed) products.

Although many AEs are detected by spontaneous reporting systems, these systems have limitations that hamper signal detection, says Van Doren. Beyond the study setting, only a small fraction of post-marketing AEs—even the most severe toxicities—are reported to drug manufacturers or regulators. Improving the ability of health care professionals to recognize and report AEs will undoubtedly be a focus of future technology. Applications for the iPad and similar devices will allow health care professionals and patients to send AE information directly to FDA. Drug manufacturers typically have product-specific or corporate websites or call centers for reporting of AEs, although as Van Doren notes, that information is often re-entered into a corporate safety database again when it gets reported to regulators. Paper-based reporting systems are so resource-intensive that drug manufacturers and regulators not using electronic AE capture systems have a backlog of hundreds to thousands of cases awaiting manual entry at any one time.

> Just staring at raw individual patient safety data, without analysis, is no way for the sponsor, investigator, or IRB to protect patients.
>
> — SALLY VAN DOREN, BioSoteria

Another concern with AERS, according to Uber, is that it gathers information from multiple sources, potentially "over-representing" information. A physician could conceivably report a SAE to both FDA and the drug manufacturer, while the same event could be reported by an attorney without necessarily being flagged as a duplicate. "To look at data at the aggregate level and not the individual case level can lead to the wrong conclusions," she says. "It all comes back to the accuracy and quality of the source data."

The FDA's remedy to under-use of its MedWatch AE reporting system is the Sentinel Initiative. Launched in May 2008, this aims to create a national electronic system for tracking reports of AEs linked to FDA-regulated products. The idea is to query diverse health care data holders—including electronic medical record (EMR) systems, administrative and insurance claims databases, and registries—to quickly evaluate potential safety issues. Data owners are tasked with mining for answers to FDA-posed questions and submitting summary results. However, it may be a few years before the Sentinel Initiative "comes good," cautions Jolley. As of December 2011, progress had been made on EMR querying methods but "not many" new safety issued had been identified.

The FDA has been instructed by Congress to mine at least 100 million EMRs—about one-third of the U.S. population—for signs of safety problems with drugs, says Paul Watkins, director of the Hamner-University of North Carolina (UNC) Institute for Drug Sciences. But without any validated tools to do this, "most safety signals that pop out will probably be false ones," he says. In the absence of randomization to treatments and placebo controls, interpreting the data is "extremely challenging." The Ob-

servational Medical Outcomes Partnership had mixed results trying to find EMR evidence of known drug troubles during beta testing of the Sentinel system, he notes.

Watkins predicts the most immediate effect of the Sentinel Initiative will be to make FDA reviewers—already spooked by the recent Vioxx and Avandia disasters—"more cautious." In the long term, once reliable approaches have been developed to screen EMRs for AEs, the Sentinel Initiative might allow a graded approval of certain drugs. One possible scenario is a new drug approved for use in health care networks where EMR data get mined in real time, allowing for rapid, real-world safety assessments.

Corporate Due Diligence

Based on audits of 50 pharma firms on three continents, SJ Pharma Consulting says there is a lot drug makers can do themselves to beef up product safety. Jolley developed a 50-page checklist of safety-related questions, best practices and regulatory guidances to help companies tease out deficiencies. Safety data exchange agreements with partners and subcontractors, for example, may be weak or absent entirely. Participants in a U.S. clinical trial were recently put at risk when a Japanese drug licensor failed to inform the sponsor of a label change indicating potential liver toxicity. The obligation to report any drug risk was never written into the safety data exchange agreement by the U.S. licensee, whom the FDA holds responsible for maintaining the integrity of the "supply chain of safety information," says Jolley.

Companies also sometimes struggle to report unexpected SAEs within 15 days, as required by FDA. Last November, it took one sponsor's safety group three months to learn that one of its trial participants had been hospitalized with a heart problem, because the reporting investigator used the wrong fax number. The SAE was entered into the EDC system, Jolley notes, but the information wasn't passed on to the safety team, so unsuspecting patients continued being enrolled.

Debate continues about how useful computers are for linking AEs to drugs, because statistically suggestive causalities ultimately require a "prepared mind" to prove, says Jolley. Data mining tools are designed for use in the post-marketing environment, where both the numerator (drug-associated AEs) and denominator (the drug's universe of users) are largely unknown. The software thus seeks clues either in public databases such as AERS or the World Health Organization's VigiBase or in proprietary big pharma databases. The denominator is approximated by comparing the number of AEs reported for a drug to the average number of drug-event combinations reported for other drugs.

Corporate desire to ensure drugs are approved, especially among small companies, may also be compromising the clinical trial enterprise, says Jolley. Sponsors tend not to use signal detection techniques to proactively identify problems. The good news is that the European Medicines Agency now compels companies to annually submit a Development Safety Update Report (DSUR) of therapies under development, much as they do for marketed drugs. FDA plans to do likewise. DSUR contains ten elements not in the IND annual report used in the U.S., including cumulative summary tabulations of all AEs of special interest and an overall safety assessment. "This should encourage companies to take a more holistic view of the

Bio-IT World Best Practices AWARDS 2012

The 2012 *Bio•IT World* Best Practices competition has released its call for entries. Since 2003, *Bio•IT World*'s Best Practices competition has been recognizing outstanding examples of technology and strategic innovation initiatives across the drug discovery enterprise.

The awards attract an elite group of life science professionals: executives, entrepreneurs, innovators, researchers and clinicians responsible for developing and implementing innovative solutions for streamlining the drug development and clinical trial process. *Bio•IT World*'s distinguished peer-review panel of judges has reviewed more than 400 entries in the program's history.

Entries will be accepted in six categories: Clinical & Health-IT; IT infrastructure/HPC; Informatics; Knowledge Management; Research & Drug Discovery; and Personalized & Translational Medicine. The 2012 winners will receive a unique crystal award to be presented at the Bio-IT World Conference & Expo in Boston, April 24-26, 2012. Winners and entrants will also be featured in *Bio•IT World*.

For more information on the program and to download the entry form, please visit www.bio-itworld.com/bestpractices.

Deadline for Entry: *Extended to* **February 15, 2012**

safety of drugs in development," Jolley says.

Hiring out safety surveillance functions to less well-qualified employees of offshore contract safety organizations is another growing problem, says Jolley. For sponsors, this is an issue of control that can be mediated by better contractual language and perhaps higher allowable fees for the advanced level of medical decision making required.

Outsourcing of safety surveillance is becoming more common for drug makers, says Uber. With limited real-world data in the early post-approval phase, companies are obliged to employ the FDA's risk evaluation and mitigation strategies (REMS) that might include restricted product distribution or creation of patient registries. Based on the sentiment of 28 organizations recently surveyed by the Tufts University Center for the Study of Drug Development, REMS may not be packing much punch: only 22% think the REMS system has improved safety. As products mature, companies can learn more from safety signaling tools. But the technology often yields a large volume of potential signals that may or may not be causally related to a drug, and can require extensive resources to investigate, says Uber.

Informed Dosing

Both FDA and National Institutes of Health (NIH) are clearly interested in improving the science behind the assessment of drug risks, and support research in priority areas of regulatory science such as adaptive trial design. The regulatory science program has recently focused on novel approaches for testing drug efficacy and safety prior to clinical trials, notably *in vitro* cell-based technologies that can mimic how human tissues respond to drugs. The Defense Advanced Research Projects Agency recently committed over $70 million to developing a "human on a chip" that can assess the safety of medical counter measures to a biological attack when there is no time for animal studies. Pre-clinical animal models are time-consuming and expensive, not to mention relatively poor predictors of results in humans. In fact, safety-related issues—and more specifically, actual or feared liver toxicity—are the main reason drugs fail, contends Watkins.

Cardiac arrhythmia disturbances, as evidenced by a prolonged QT interval, were until recently the most frequent culprit in drug recalls. But early-phase EKG testing on healthy volunteers has become standard practice, making liver meltdown the major unresolved safety issue. In the absence of a regulatory path for drug-induced liver injury (DILI), FDA sometimes requires expanded clinical trials to improve the odds of detecting a latent liver issue. The "disaster" is not just the time and costs of these extended trials, but potentially billions of dollars for the years of lost patent life "we all end up paying for."

Watkins is chairman of the steering committee for the NIH-funded Drug-Induced Liver Injury Network, which since 2004 has been endeavoring to genetically characterize people who experience uncommonly severe liver reactions to a drug. Separately, the Hamner-UNC Institute has been attempting to create a computer model of DILI to make animal models more predictive—and ultimately help eliminate animals from drug development entirely. The European Union is taking the lead in this arena with its Innovative Medicines Initiative (IMI) that has set aside 2 billion euros ($2.6b) for matching grants going to pharma companies engaged in regulatory sci-

ence research. The IMI's goal, says Watkins, is to "make Europe the worldwide home for pharmaceutical research."

Meanwhile, the Hamner-UNC Institute is about to launch a DILI-sim initiative with partners AstraZeneca, GlaxoSmithKline, Novartis, and other pharma companies, says Watkins. The goal is to model significant inter-species differences and susceptibilities to liver toxicity, to help inform first-in-man dosing. Watkins argues that part of the problem in achieving efficacy is the failure to administer a sufficiently high dose. Currently, drugs are typically given in doses "at least ten times lower than what causes liver toxicity in animals," he says.

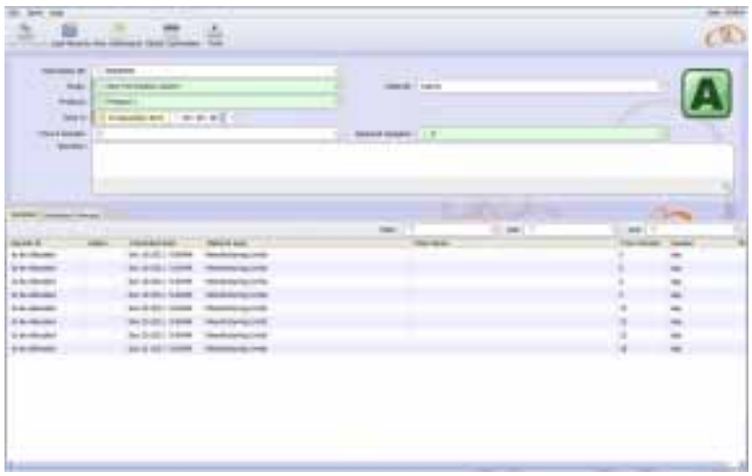
> Better biomarkers of organ toxicities would be a huge advance for trial safety.
>
> — **PAUL WATKINS,** Hamner-UNC Institute for Drug Sciences

Corporate partners are providing financial and in-kind resources as well as data from compounds that looked promising in animals but subsequently fail in humans. The FDA, which intends to apply lessons from DILI-sim in regulatory decisions, is also supportive. The first iteration of the computer model will be distributed to partners early in 2012.

Previous research from the Hamner-UNC Institute suggests a simple urine test could predict which clinical trial participants are at risk for DILI, based on specific metabolite patterns predictive of mild liver damage associated with acetaminophen. Better biomarkers of organ toxicities would be a huge advance for trial safety, says Watkins. The four blood tests currently available for detecting liver problems have been in use for 50 years, produce false-positive results, and miss sudden-death liver injury that can inexplicably arise weeks or months after ingestion of a drug.

"Companies do not routinely collect and keep blood and urine samples from clinical trials," says Watkins. "If they do, there's no standard way to handle them and link [the samples] to patient data." Watkins is leading the push to create a standardized safety database linked to a biospecimen repository. In the cardiovascular arena, FDA already prescribes how EKGs should be performed and the data warehoused. The mandate helped spur formation of the Cardiac Safety Research Consortium for ongoing evaluation of medical products. Watkins would like to launch a similar consortium, initially focused on liver safety, at the Hamner-UNC Institute.

However, increasingly robust patient privacy requirements could make creation of a biospecimen repository problematic, warns Uber. "If a study sponsor retains samples, it has to inform study participants. And if a sponsor doesn't know how those samples may eventually be analyzed and evaluated, they will need to obtain permission for 'research not yet specified.' Language that provides the sponsor unlimited future access to retained biological specimens is a permission patients are likely to question, and may be reluctant to grant."



Stability Trial Software

Two Fold Software is adding a user-friendly Stability Trials software module to Version 1.1 of the company's innovative Qualoupe LIMS software. The new Stability Trials software has been specially tailored for organizations that need to run stability trials over any given period of time and different storage conditions. Using the software the user can create protocols for parameters such as sampling regimes, detailing storage conditions, sampling frequency and time units. When running a trial the user can select the protocol which would then create the scheduled samples for the trial. The material on trial would associate the methods to be performed on the samples. Trials can be associated with studies and projects. The Stability Trials software also enables the LIMS to provide prompts for sample pull schedules and the storage location.

Product: Stability Trials module for Qualoupe LIMS
Company: Two Fold Software
For more information: www.twofold-software.com

Gene Design in Open Language

DNA2.0 has announced the integration of the Synthetic Biology Open Language (SBOL) into the company's groundbreaking gene design and assembly tool, Gene Designer. SBOL is a synthetic biology standard exchange format that allows scientists to share, store, explore and publish genetic elements created through computer design. SBOL is the first standardized information exchange framework created specifically for bioengineering. This newly-released language forms the basis for Gene Designer's integration of genetic parts from BIOFAB and enables a new feature in the application called Diagram. Diagram provides a clean visualization interface of designed sequences and eases scientific collaboration by enabling graphic exports of sequences.

Product: Diagram visualization interface for Gene Designer
Company: DNA2.0
For more information: www.DNA20.com

Expanding Genomics Suite

Partek has released two new products: Partek Flow and Partek Pathway. Bookending the Partek Genomics Suite, Partek Flow provides upstream alignment, QA/QC and quantification of next generation sequencing (NGS) data, and can be installed on a desktop, cluster, or cloud environment. Partek Pathway provides contextualization of the gene relationships identified within Partek Genomics Suite, allowing users to study gene-gene and gene-protein relationships with the ability to identify and display significantly affected pathways, search for specific pathways and genes, and color code them based on their p-values and fold changes. Together, the tools offer a comprehensive start-to-finish solution for NGS data analysis providing alignment, QA/QC, exploratory analysis, statistical analysis, interactive visualization, genomic integration, and biological interpretation. They support raw data from all major commercial NGS and microarray technologies and the five core assays: RNA, DNA, microRNA, ChIP, and Methylation.

Products: Flow and Pathway
Company: Partek
For more information: www.partek.com

Kinase Services

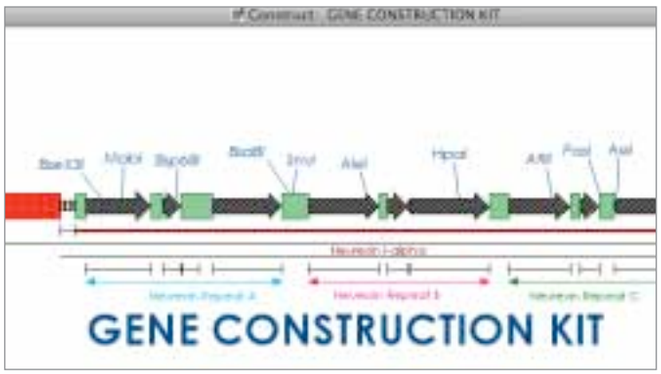
Caliper Life Sciences has launched In-Cell KinaseScreen, a new services offering based on a panel of cellular assays that measure the effects of drug candidates on the functional activity of kinases involved in critical cell-signaling pathways. The technology platform and initial panel of nine assays were developed in collaboration with Pfizer. The service is immediately available for compound screening, profiling, and custom assay development services for Caliper Discovery Alliances and Services' (CDAS) clients, ranging from academia to pharmaceutical companies.

Product: In-Cell KinaseScreen
Company: Caliper Life Sciences
For more information:
www.caliperLS.com/CDAS

Next Gen Analysis Server

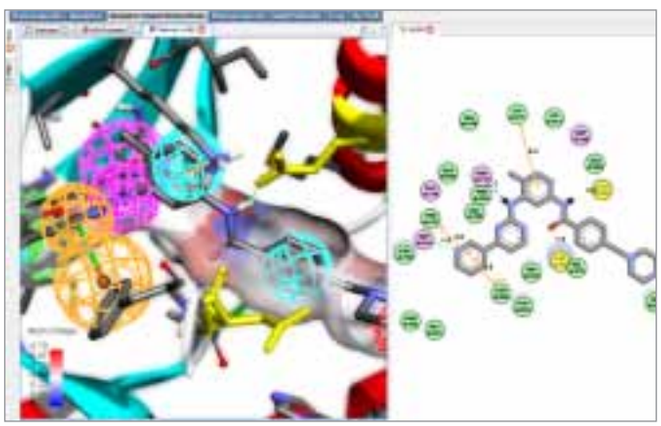
Biomatters has released version 1.5 of Geneious Server, an NGS analysis product that connects biologists to powerful Linux-only software algorithms and high-performance hardware resources such as clusters. Geneious Server 1.5 includes the ability to connect to clusters using MOAB/Torque in addition to Oracle's SGE and Platform Computing's LSF, adds support for Queue Licenses and features TopHat for mapping millions of RNA-Seq reads against a reference genome in minutes.

Product: Geneious Server v.1.5
Company: Biomatters
For more information:
www.geneious.com

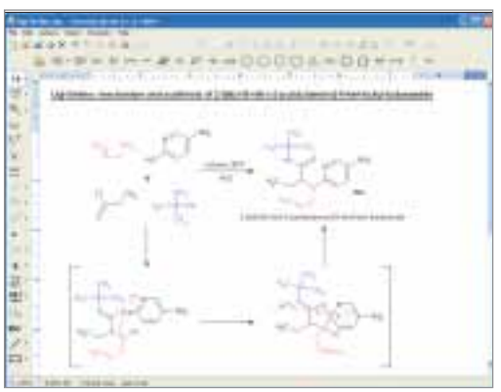


Company: Textco BioSoftware
Product: Gene Construction Kit (GCK)
Description: Download GCK to see for yourself how this flexible plasmid mapping tool makes analyzing and annotating DNA cloning projects faster and easier—making your lab more productive.
Try if for free: http://bit.ly/FreeTextco

Company: ARX
Product: CoSign Digital Signatures
Description: CoSign is the most widely-deployed digital signature solution in the Life Sciences industry, employed by over 20,000 FDA-regulated organizations.
Try it for Free: http://bit.ly/FreeARX



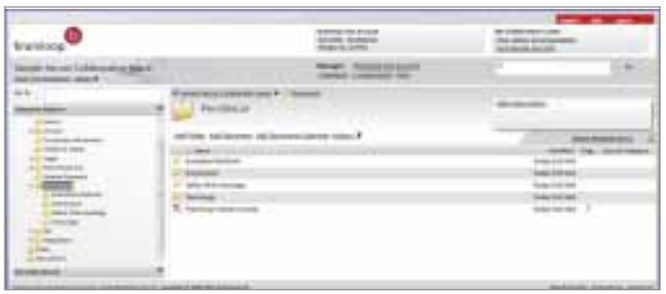
Company: Accelrys
Product: DS Visualizer
Description: DS Visualizer is a commercial-grade graphics visualization tool for viewing, sharing, and analyzing protein and modeling data.
Try it for Free: http://bit.ly/FreeVisualizer



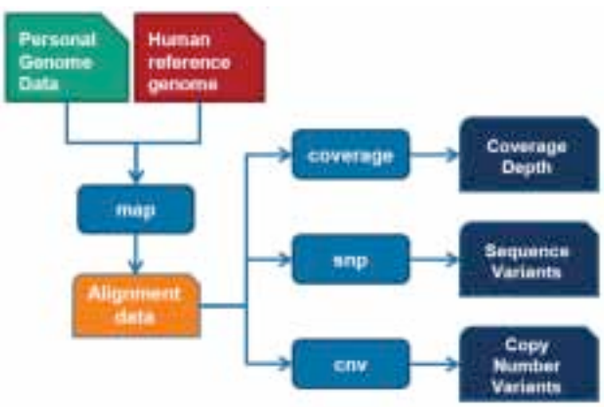
Company: Accelrys
Product: Draw
Description: Accelrys Draw enables scientists to draw and edit complex molecules and chemical reactions with ease, facilitating the collaborative searching, viewing, communicating and archiving of scientific information.
Try it for Free: http://bit.ly/FreeAccelrysDraw

Feature Your Free Download To have your free trial or download listed on this page, contact Allison Proffitt, Managing Editor, at aproffitt@healthtech.com

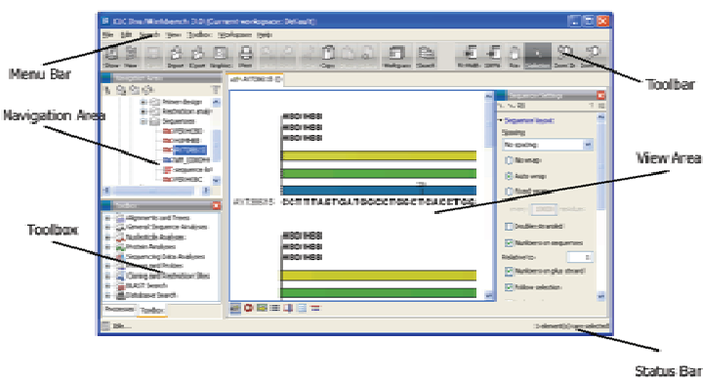
Company: Denodo Technologies
Products: Denodo Platform
Description: Denodo Data Virtualization abstracts data from enterprise, Web/Cloud and unstructured data sources and combines it into a "virtual" data layer that fulfills critical information needs, faster with fewer resources. Life sciences companies use Denodo to deliver real-time data services to power research, clinical, compliance and sales initiatives.
Try it for free: http://bit.ly/FreeDenodo



Company: Brainloop
Product: Brainloop Secure Dataroom
Description: Brainloop offers a highly secure online workspace for drug partnering plus collaboration on clinical trials, regulatory filings, manufacturing. Fast setup and prevent printing, forwarding, saving.
Try it for Free: http://bit.ly/FreeBrainloop



Company: Real Time Genomics
Product: RTG Investigator
Description: RTG Investigator is advanced sequence analysis software for variant detection and metagenomic analysis with large-scale Illumina, Complete Genomics and Roche 454 data sets.
Try it for free: http://bit.ly/FreeRTG



Company: CLC bio
Products: CLC Genomics Workbench
Description: CLC Genomics Workbench analyzes high-throughput sequencing data faster than it is produced by implementing a SIMD-accelerated assembly algorithm.
Try it for Free: http://bit.ly/FreeCLCbio

Keep abreast of the variety of educational events in the life science industry that will help you with your business and professional needs. To preview a more in-depth listing of educational offerings, visit the "Events" section of **bio-itworld.com**.

To list an educational event, contact Lisa Scimemi at **lscimemi@healthtech.com**.

CHI Events

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

Summit for Clinical Operations Executives (SCOPE)
February 7-9, 2012 | Miami, FL

Cloud Computing
February 19-20, 2012 | San Francisco, CA

Next Generation Pathology
February 19-20, 2012 | San Francisco, CA

Molecular Med TRI·CON
February 21-23, 2012 | San Francisco, CA

Integrated R&D Informatics for Knowledge Management
February 21-23, 2012 | San Francisco, CA

Bioinformatics & Cancerinformatics
February 21-23, 2012 | San Francisco, CA

X-GEN CONGRESS & EXPO
March 5-8, 2012 | San Diego, CA

microRNA in Human Disease and Development
March 12-13, 2012 | Cambridge, MA

Bio·IT World CONFERENCE & EXPO '12
April 24-26, 2012 | Boston, MA

PEGS: the essential protein engineering summit
April 30 - May 4, 2012 | Boston, MA

Strategic Alliance Management Congress
May 7-9, 2012 | Philadelphia, PA

Biomarker World Congress
May 21-23, 2012 | Philadelphia, PA

WPC World Pharma Congress
June 5-7, 2012 | Philadelphia, PA

Bio·IT World ASIA CONFERENCE 2012
June 6-8, 2012 | Singapore

Visit **BarnettInternational.com** for detailed information on Barnett's live seminars, interactive web seminars, on-site training programs, customized eLearning development services and publications.

Barnett Live Seminars

Auditing Techniques for Clinical Research Professionals
February 6-7, 2012 | Boston, MA

Data Management in the Electronic Data Capture Arena
February 6-7, 2012 | Boston, MA

Comprehensive Monitoring for Medical Devices
February 7-9, 2012 | Philadelphia, PA

CRA & CRC Beginner
February 7-9, 2012 | Philadelphia, PA

Clinical Project Management: Advanced
February 9-10, 2012 | Boston, MA

Clinical Project Management: Introduction to Practical Clinical Trial Planning for Project Managers
February 14-15, 2012 | Philadelphia, PA

Monitoring Clinical Drug Studies: Intermediate
February 16-17, 2012 | Philadelphia, PA

Conducting Clinical Trials Under ICH GCP
February 16-17, 2012 | San Diego, CA

Regulatory Intelligence
February 17, 2012 | San Diego, CA

Monitoring Clinical Drug Studies: Beginner
February 20-22, 2012 | Philadelphia, PA

Adverse Events: Managing and Reporting for Medical Devices
February 23-24, 2012 | Philadelphia, PA

Advanced Good Clinical Practice
February 23-24, 2012 | Philadelphia, PA

Medical Device Approval Process
February 27-28, 2012 | San Diego, CA

The Highly Effective CRA
February 27-28, 2012 | San Diego, CA

Developing Clinical Study Budgets
March 1, 2012 | San Diego, CA

Good Clinical Practice for Laboratory Scientist
March 1, 2012 | San Diego, CA

Pharmacovigilance Audit
March 1, 2012 | San Diego, CA

Adverse Events: Managing and Reporting for Pharmaceuticals
March 1-2, 2012 | Philadelphia, PA

Clinical Project Management: Intermediate
March 1-2, 2012 | Philadelphia, PA

Conducting Clinical Trials in Resource Limited Settings
March 1-2, 2012 | Boston, MA

Patient Recruitment and Retention
March 1-2, 2012 | Boston, MA

Introduction to FDA
March 8-9, 2012 | Philadelphia, PA

Barnett Web Seminars

Corrective Action Plans: Essential Documentation of a Site's Response to GCP Deficiencies
January 23, 2012

Adverse Events for Medical Devices
January 23, 2012

Comparing FDA and Health Canada Regulations: Using an ICH GCP Framework
January 26, 2012

Critical Decision Points in Design & Conduct of Patient Registries
January 26, 2012

New FDA Guideline to Informed Consent
January 27, 2012

March 2011 Compliance Program Guidance Manual (CPGM) Update: Recommendations for the Inspection of Sponsors, CROs, and Monitors
January 30, 2012

RECIST 1.0 and 1.1: Overview and Data Challenges in Oncology Clinical Trials
January 30, 2012

Train-the-Trainer: Successful Web-Based Training Strategies
January 31, 2012

NEW! Investigator Selection Criteria and Strategies for Investigator Qualification
January 31, 2012

Visit **bio-itworld.com** to browse our extensive list of complimentary Life Science white papers, podcasts and webcasts.

To learn more about developing a multimedia lead generating solution, contact Lisa Scimemi at **lscimemi@healthtech.com**

Whitepaper

Accelerating Decision Making Processes in Life Sciences R&D

Sponsored by HP and Microsoft

Life sciences organizations need an easy way to access and share information to make timely R&D decisions. Read this white paper to learn:

- How IT and business challenges complicate decision-making efforts
- How to avoid missed opportunities, duplicate work, and inefficient use of time
- How a jointly developed HP and Microsoft business intelligence appliance provides a way for life sciences researchers and managers to accelerate their analysis and decision-making work.

Download Today

Webcast

Meeting Compliance Needs through Enhanced Management of Patient Data

Sponsored by: Dell Compellent

Healthcare providers understand the pressure to keep the data center highly-available at all times. These organizations consistently try to manage rapidly growing digital patient data, while adhering to regulatory compliance that dictates how long it must be stored. Learn how Radiology Ltd. has utilized the Dell Compellent SAN to improve the performance and scalability of its data center, while gaining critical transparent reporting to help enable better patient care.

Presenter: Colin Schmugge, Data Center Manager at Radiology, Ltd.
Moderator: Dr. Kevin Davies, Editor-in-Chief, *Bio-IT World*

Download Now

Podcast

Meeting Today's Challenges in Clinical Trial Supply Management

Sponsored by: Medidata Solutions Worldwide

Setting up and managing the clinical trial involves many complex procedures. Among the most challenging are planning and executing the logistics of the trial's clinical supplies. This podcast focuses in depth on the following topics which trace current practices and future evolution of this crucial aspect of clinical trials:

- Current practices in clinical trial logistics
- Comparing advances in clinical supply practices to other aspects of clinical trials
- Where current practices fall short of meeting the challenges
- Trends and evolving improvements that may change the way logistics are conducted

Listen Now

Web Symposium

Whole Genome Sequencing (WGS): The Next Diagnostics Platform

April 3, 2012

healthtech.com/zmdsq

RSA® CONFERENCE 2012

FEBRUARY 27-MARCH 2 | MOSCONE CENTER | SAN FRANCISCO

WE STAND UNITED.
JOIN US AT RSA® CONFERENCE

THE GREAT CIPHER MIGHTIER THAN THE SWORD

SAVE $400
Register before Friday, January 27!

Today's hacktivists publicize sensitive data, bring down networks, and cut off services. And the collateral damage to governments, businesses and consumers is mounting. To defend against this latest onslaught of cybercrime, our community must join forces. At RSA® Conference 2012, you will unite with like-minded security professionals and pundits, exchange useful and effective strategies and become stronger in the face of these growing security threats.

RSA Conference 2012 is the premier event where you will find the insights and resources you need to stop hacktivist attacks and protect your domain.

DEEPEN YOUR KNOWLEDGE
Attend over 220 expert-led sessions.

SHARE YOUR EXPERIENCE
Network with prominent security visionaries.

DISCOVER YOUR OPTIONS
Explore innovative products from over 350 exhibitors.

ARM YOUR ORGANIZATION
Gain practical, real-time solutions to implement.

REGISTER NOW!
www.rsaconference.com/bioit

Platinum Sponsors: ca technologies, Cisco, hp, McAfee

Gold Sponsors: ActivIdentity, nCircle, SOPHOS, IT Security Made in Germany, websense

Silver Sponsors: Fallen Vault, axway, BlueCoat, Check Point, Dell SecureWorks, eset Internet Security, M86 Security, NARUS, Safelight, Secunia, TITUS, TREND MICRO, tripwire, Trustwave, verizon

Coming in *Bio•IT World*'s February Issue

- **Cancer Connection:** An in-depth profile of Foundation Medicine and CEO Michael Pellini in advance of the launch of the company's state-of-the-art NGS cancer test that will sequence some 200 genes from cancer samples, providing physicians with detailed information on mutation status, potential drug candidates and clinical trials.
- **Data Orchestration:** How IDBS is deploying workflow solutions and electronic lab notebooks to aid data management and advance translational medicine.
- **Is IT a Cloud?** In a further sign of growing acceptance of cloud computing, a biopharma consulting firm describes how and why it put its entire IT infrastructure into the cloud.
- **Plus:** One big pharma's approach to **medical imaging**, supercomputing prediction of **adverse events**, and a round-up of news and **trends in ELNs**.

Contact Us

EDITOR-IN-CHIEF
Kevin Davies (781) 972-1341
kdavies@healthtech.com

MANAGING EDITOR
Allison Proffitt
(617) 233-8280
aproffitt@healthtech.com

ART DIRECTOR
Mark Gabrenya
(781) 972-1349
mgabrenya@healthtech.com

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

VP SALES — ACCOUNTS A–K
Alan El Faye
(213) 300-3886
aelfaye@healthtech.com

ACCOUNT MANAGER — ACCOUNTS L–Z
Tim McLucas
(781) 972-1342
tmclucas@healthtech.com

CORPORATE MARKETING COMMUNICATIONS DIRECTOR
Lisa Scimemi
(781) 972-5446
lscimemi@healthtech.com

WEB ADVERTISING OPERATIONS MANAGER
Franscena Schandelmayer-Davis
(781) 972-1351
fsdavis@healthtech.com

ADVERTISING OPERATIONS COORDINATOR
Stephanie Cline
(781) 972-5465
scline@healthtech.com

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

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

Cambridge Healthtech Institute
PRESIDENT
Phillips Kuhl

Contact Us:
editor@healthtech.com
250 First Avenue, Suite 300
Needham, MA 02494

CONNECT WITH US!

Companies in This Issue

C CONTENTS

Advertiser Index

Bio·IT World

CAMBRIDGE HEALTHTECH MEDIA GROUP

Indispensable Technologies Driving Discovery, Development, and Clinical Trials

VOLUME 11, NO. 1

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

Bio·IT World is published 10 times a year by Cambridge Bio Collaborative, 250 First Avenue, Suite 300, Needham, MA 02494. *Bio·IT World* is free to qualified life science professionals. 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).

Subscriptions: Address inquires to Bio-IT World, 250 First Avenue, Suite 300, Needham, MA 02494 888-999-6288 or e-mail kfinnell@healthtech.com

What is (Quantitative) Systems Pharmacology?

By John Russell

Just a few days apart in the middle of October, Harvard Medical School (HMS) announced a broad initiative in systems pharmacology and NIH released a like-minded white paper, *Quantitative and Systems Pharmacology in the Post-genomic Era: New Approaches to Discovering Drugs and Understanding Therapeutic Mechanisms*. That's a fair amount of attention from two very big guns on a topic that may ring both familiar and unfamiliar.

Actually systems pharmacology has been loosely percolating for some time. While not exactly a splinter from the systems biology world, systems pharmacology shares much of systems biology's methodology and conceptual framework. Indeed, the Harvard Medical School systems pharmacology initiative is being launched from its systems biology department and one of the NIH paper's lead authors (Peter Sorger) is a member of the department.

Nuanced definitional wrangling aside (sure to be fun), what distinguishes systems pharmacology is its laser-like focus on compounds and how they perturb biological systems and pathways. How specifically do compounds—failed and successful drugs as well as others—work in the body? What are the detailed mechanisms? How are they influenced by various 'omics? How do they vary by tissue? etc. This is a great idea, which is not to say such activities weren't part of the broad the systems biology world.

The practical implications of such a compound-centric approach are exciting: new targets, new screens, new markers, new understanding of drug failure mechanisms. Indeed sophisticated drug failure analysis may be one of SP's most promising goals and eventually most rewarding contributions. (FDA should open its treasure trove of information for these efforts, but that's another matter). It's a fairly firm rule-of-thumb that the biopharmaceutical industry doesn't spend a lot of time or resources investigating why drugs fail.

Helping launch, promote and develop systems pharmacology as well as delivering concrete SP results are the main goals of the Harvard initiative. It is being led by William Chin, executive dean for research at HMS; Marc Kirschner, chairman of the HMS Department of Systems Biology; Peter Sorger; and Tim Mitchison, deputy chair of the department. The ambitious plan calls for adding

faculty (10), tackling research, establishing new research infrastructure, and collaborative outreach to biopharma, other academia, and biomedical communities.

SP in Practice

According to the Harvard SP announcement, "The initiative will support both new approaches in translational science, such as failure analysis on unsuccessful drugs and use of chemical biology to develop probes of biological pathways. It will also include a new educational program, one that develops a new generation of students, postdoctoral fellows and physician-scientists, the future leaders in academic and industrial efforts in systems pharmacology and therapeutic discovery."

Marc Kirschner

Obviously everything won't happen at once, but Kirschner and colleagues are charting the path forward. "One project already going on in our place and at MGH (Massachusetts General Hospital) is a study of the drug Taxol," he says. "It delivers nice understandable effects in some cell cultures but the effect on human xenograph tumors in the mouse, when viewed by intravital imaging, shows the drug acts in a very different way there. We have a lot to understand about drug action not only in cells but also in more complex environments."

Kirschner resists the notion that systems pharmacology is an applied science, "I'm not a big believer in distinguishing at least in the early stage applied from fundamental. People were trying to understand how to make steam engines more efficient and we ended up with thermodynamics. [Y]our goal may be to make nylon but you end up having to figure out all physics and chemistry surrounding polymers and how they assemble and interact. [That] gets to be very fundamental."

Biopharma's response has been positive, Kirschner says. Industry execs he's talked to believe SP can have practical benefits. They also agree the effort is best led by academia, given that economic pressures are prompting even the biggest drug makers to downscale research, especially basic research. It's also more difficult for biopharma to attract the multi-discipline talent required to pursue SP properly, says Kirschner.

One question is how to measure SP progress and impact. Broadly speaking, "[It] can be looked at in two different ways," says Kirschner. "If the goal was to achieve some specific target for a specific disease, I think the likelihood of that happening quickly is small. On the other hand if the goal is to achieve significant information and tools useful for the development of targets in some disease, that likelihood is quite high."

In the end, he says, systems pharmacology is an open-ended exercise and biopharma will have access to the results. Systems pharmacology and framing it as a compound-driven exercise is great idea. Let's see where it leads.

Cambridge Healthtech Institute's Inaugural

Bio-IT World ASIA CONFERENCE 2012

June 5-8, 2012 | Marina Bay Sands, Singapore

FOCUSED SESSION TRACKS

- IT Infrastructure and the Cloud
- Next-Generation Sequencing Data Management and Interpretation
- Bioinformatics
- Drug Discovery Informatics

MARK YOUR CALENDARS

Recent advances in scientific technology have left us confronted with the task of discovering scientific knowledge from enormous amounts of data generated in genomics, pharmaceutics, medicine and other life science areas.

Cambridge Healthtech Institute's Inaugural **Bio-IT World Asia Conference**, building upon the exciting momentum of the flagship **Bio-IT World** conference in Boston, will provide an ideal occasion for both the life science community and the information technology industry to meet and discuss the challenges and solutions on data management infrastructure, interoperability, and the complexity of data analysis in biomedical research and drug development process.

Official Publication

Bio-IT World Bio-ITWorldAsia.com

ADVISORY BOARD

- M. K. Bhan, M.B.B.S, M.D., D. Sc., Secretary, Government of India, Department of Biotechnology, Ministry of Science & Technology
- Chris Blessington, Senior Director, Marketing & Communications, Isilon
- Han Cao, Ph.D., Founder, CSO, BioNanomatrix, Inc.
- Laurie Goodman, Ph.D., Editor-in-Chief, (Giga)n Science, BGI-Shenzhen
- Sean Grimmond, Ph.D., Professor, Molecular Bioscience, University of Queensland
- Yike Guo, Ph.D., Professor in Computing Science, Imperial College London
- Linh Hoang, Director of Genomic Medicine, Life Technologies
- Krishan Kalra, Chairman & CEO, BioGenex Labs Inc.
- Peter Little, Ph.D., Research Director, Life Science Institute, National University of Singapore
- Yusuke Nakamura, Ph.D., Professor, Laboratory of Molecular Medicine, Institute of Medical Science, The University of Tokyo; Secretary General, Office of Medical Innovation, Cabinet Secretariat, Government of Japan
- Pauline Ng, Ph.D., Group Leader, Computational and Mathematical Biology, Genome Institute of Singapore
- Stephen Rudd, Ph.D., CSO, Malaysian Genomics Resource Centre Berhad
- Parthiban Srinivasan, Ph.D., President & CEO, Parthys Reverse Informatics, India
- Tin-Wee Tan, Ph.D., Deputy Head of Department of Biochemistry , National University of Singapore
- Alain Van Gool, Ph.D., Head Molecular Profiling, Translational Medicine Research Center, Merck Sharpe & Dohme

Cambridge Healthtech Institute
CHI 250 First Avenue, Suite 300 | Needham, MA 02494 | T: 781-972-5400 or Toll-free in the U.S. 888-999-6288 | F: 781-972-5425