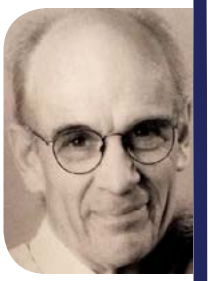
ISCB
INTERNATIONAL SOCIETY FOR COMPUTATIONAL BIOLOGY

Newsletter

SPRING 2015
VOLUME 18
ISSUE 01

Features

FASEB Names **David Rocke** Treasurer-elect



Peter H. Sellers, among the earliest researchers on DNA and protein sequence comparison, died of cancer on November 15, 2014, at the age of 84.

In this issue:

A global community dedicated to accelerating scientific discovery through computation.

Valencia Begins Term as ISCB President

by Christiana N. Fogg, Ph.D.

Alfonso Valencia officially began his term as ISCB President on January 21, 2015, succeeding Burkhard Rost of the Technical University of Munich. Valencia is the Vice Director of Basic Research at the Spanish National Research Centre (CNIO) in Madrid, Spain, where he also directs the Spanish National Bioinformatics Institute and leads the Structural and Computational Biology group.

Valencia studied population genetics and biophysics at the Complutense University of Madrid and received his PhD in molecular biology from the Autonomous University of Madrid in 1988. He did his postdoctoral training in the lab of Chris Sander at the European Molecular Biology Lab (EMBL) in Heidelberg, Germany, where he became acquainted with computational approaches to biological problems and developed methods to understand and predict protein functions.

Following his postdoctoral training, Valencia formed the Protein Design Group at the Spanish National Center for Biotechnology (CNB) in the CNIO, which has evolved into the Structural Computational Biology Group. Valencia's primary research interest is to understand the molecular bases of cancer progression, with a particular focus on the interplay between genomics and epigenomics during the course of cancer development. He and his group have developed novel software platforms to extract, integrate, and represent cancer data.

Valencia sees great promise for breakthroughs in cancer research with the confluence of "omics" projects and clinical data. He said, "My work, like many of my ISCB colleagues, combines the development of original methods—something absolutely necessary in a still new research area, and collaborating with clinical groups in the implementation of robust analysis strategies." He acknowledges that he and his

computational biology colleagues are faced with some unprecedented challenges that he describes as "a complex scenario that covers the full range, from the organization of basic bioinformatics infrastructures for the analysis of the complex genomic and biomedical data, to the direct analysis of specific biomedical cases, and the research developments triggered by new scientific problems."

Valencia considers ISCB a vital organization that brings together and supports the computational biology community, especially at a time when members of the community are experiencing unprecedented demand for their unique expertise. Valencia has long been involved in building and developing ISCB; he has served as a founding Board member, Vice President, Outreach Chair, Awards Chair, Fellows Chair, and was selected as an ISCB Fellow in 2010.

As President, he wants to build on the healthy growth of ISCB by focusing on three areas:

- To make ISCB not only a successful scientific society but also a society able to represent the professional interests of our members, reflecting the increasing role of bioinformatics and computational biology in industry
- To improve relations with our affiliated groups by providing them support for their activities
- To provide support for high quality activities, in particular those promoted by our junior PIs. Such activities with intrinsic value will contribute to the positive perception of the science and technology developed by ISCB

Precision medicine is a term coined recently to describe tailored medical care for individuals based on genetic or molecular diagnostic information. Valencia sees ISCB members as vital leaders and players in this field. He said, "ISCB members are already an intrinsic part of the development of precision medicine, a discipline that cannot be understood without the associated computational methods and bioinformatics resources. These scientific and technical contributions are recognized and publicized in our meetings and publications. Precision medicine has been, and still is, one of the central topics in the ISCB meetings and activities."

Valencia's election as President is a reflection of the esteem his colleagues hold for him. Anna Tramontano, co-chair of the ISCB Nominating Committee that worked to select Valencia as a candidate, said, "Alfonso Valencia has an exceptional portfolio of scholarship and leadership, and, just as importantly, commitment to the Society. He has been charting new frontiers in our science with vision and dedication, and brings great academic achievement, proven leadership skills, integrity, commitment to our mission and values. I am convinced that the breadth and depth of his experience, his vision and, last but not least, his personality, will enable the Society to preserve its heritage and to meet the challenges of the future at the highest level."

Indeed, ISCB is poised to have a bright and fruitful future with Valencia at the helm.

ISMB/ECCB Rebooted: 2015 Brings Major Update to the Conference Program

by Christiana N. Fogg, Ph.D.

The 23rd Annual International Conference on Intelligent Systems for Molecular Biology and the 14th European Conference on Computational Biology (ISMB/ECCB 2015) is shaping up to be a phenomenal meeting with world-renowned keynote speakers and changes in the conference format that aim to make for a more streamlined and user-friendly conference. The joint ISMB/ECCB conference is held biennially and is the flagship meeting of the International Society for Computational Biology (ISCB). The meeting will take place at the Convention Center Dublin, Ireland from July 10-16, 2015 and will bring together scientists from across the globe working in a broad range of computational biology-related disciplines including genomics, structural biology, proteomics, data mining, machine learning, and systems biology.

In the past, presentations at the meeting have been organized according to tracks: Proceedings, Highlights, and Late Breaking Research tracks. Combined with the multiple track presentation, this caused frustration for attendees when choosing which sessions to attend. In 2015 all oral presentations will be presented in broad theme areas. As in the past, submissions accepted into the highly selective Proceedings track will be published in a special ISMB/ECCB Proceedings issue of Bioinformatics.

Conference co-chairs Alex Bateman, Janet Kelso, and Desmond Higgins and the conference committee undertook the reorganization effort and came up with five theme areas: **Genes, Proteins, Systems, Disease, and Data**. Batemen said, "The idea of themes is an obvious way to organize the talks. But, selecting a small number of themes that represented all computational biology was challenging. Of course many talks will potentially fit across several themes. Time will tell whether these need any tweaking for future meetings."

Kelso believes the new organization will benefit attendees and said, "We hope that organizing the meeting more thematically will mean that attendees have an easier time identifying sessions that are relevant and interesting to them."

Five leading scientists have been named as theme chairs and will organize the selection of presentations from each traditional submission category for each theme area. They are Yana Bromberg of Rutgers University (Disease), Phil Bourne of NIH (Data), Nicolas Le Novere of the Babraham Institute (Systems), Martin Vingron of the Max Planck Institute for Molecular Genetics (Genes), and Ioannis Xenarios of the University of Lausanne (Proteins).

As in the past, the keynote speaker line up features world-class scientists. The speakers include Amos Bairoch of the Swiss Institute of Bioinformatics (ISCB Fellows Keynote), Cyrus Chothia of the MRC Laboratory of Molecular Biology (Senior Scientist Award winner), Eileen Furlong of the European Molecular Biology Laboratory, Curtis Huttenhower of the Harvard T.H. Chan School of Public Health (Overton Prize winner), 2013 Nobel Laureate Michael Levitt of Stanford University, and Kenneth Wolfe of University College Dublin.

The main conference program takes place from Sunday, July 12 through Tuesday, July 14. In the tradition of previous ISMB/ECCB meetings, Special Interest Group (SIGs) and Satellite Meetings will occur prior to the conference on Thursday, July 10 and Friday, July 11. This year, the tutorials and workshops that usually precede the main meeting are being replaced by applied knowledge exchange sessions (AKES). The AKES are scheduled for Saturday, July 11 and are designed to provide interactive educational and knowledge exchange opportunities for attendees. AKES will also provide a chance for junior principle investigators to meet and exchange career advice. Six AKES have been scheduled:

Satellite Meetings

3Dsig: Structural Bioinformatics & Computational Biophysics
CAMDA 2015 Critical Assessment of Massive Data Analysis

Special Interest Groups (SIGS)

Automated Function Prediction (AFP-SIG)
Bio-Ontologies
BioVis SIG
BOSC: The 16th Annual Bioinformatics Open Source Conference
HitSEQ: High Throughput Sequencing Algorithms & Applications
Integrative RNA Biology SIG
NetBio SIG
Regulatory Genomics Special Interest Group (RegGenSIG)
VarI SIG

Applied Knowledge Exchange Sessions (AKES; previously known as tutorials workshops)

AKES 01: *Applied Knowledge Building networks for translation: how to use DREAM challenges and the Synapse platform as a research strategy*
This session will explain the rationale behind running a DREAM Challenge, the steps involved from both sides of the process (organizers and participants), the lessons learned, and the potential uses of DREAM Challenges in education. Hands-on exercises with the Synapse platform will be featured.

AKES 02: *Cytoscape 3 App Development: Variations on a theme -- "Hello World"*
Cytoscape's real power lies in the ecosystem of community-developed apps. The most common types of apps provide access to third-party biological databases, customize data import for domain-specific data sets, and perform custom analyses and workflows. During this workshop, we will demonstrate how to develop apps for Cytoscape, targeting individuals who want to take advantage of the network visualization and analysis capabilities of Cytoscape and extend it for custom use cases.

AKES 03: *Bioinformatics software testing and quality assurance*
In this workshop, we aim to create an environment to engage bioinformatics software developers, managers of bioinformatics/genomic core facilities, researchers in reliability engineering and statisticians. We hope that cross pollination of these fields would yield interesting new insight and ideas that would open new research avenue related to bioinformatics quality assurance.

AKES 04: *Open-access, cloud-based, individual-level clinical trials data - sharing, dissemination and analyses*
In this workshop, we will discuss the reasons for data sharing and the issues surrounding it, the various ways sharing is implemented, and showcase our experience in re-analysis of clinical trials data using open immunology studies data. We will also highlight our initial work on defining a minimum information guideline for clinical trials data release.

AKES 05: U*sing biological cyberinfrastructure to scale science and people – Applications in data storage, HPC, cloud analysis, and bioinformatics training*
Cyberinfrastructure (CI) is a powerful enabler for data-intensive biology. Although much investigation originates in organism-centered communities there are unifying similarities across types of datasets, algorithms, and overall goals. This workshop demonstrates how CI originally developed for the U.S. plant science community (via. the iPlant Collaborative project) serves all life sciences (animals, plants, microbes, etc.) by allowing communities to leverage pre-built CI solutions and develop application-specific components to a customized endpoint.

AKES 06: *How to navigate a bioinformatics career path*
Working with the ISCB student council, the Junior PI group and COBE COSI, this session is focused on career development. Here, we focus on training and preparing the bioinformatics professional to successfully launch and build a career. This session will be of interest to a wide audience. For students and junior PIs, you will participate in topics and practical sessions that can help build your career. For senior faculty and other professionals, it will provide ways in which you can improve skills to mentor your students and professionals building their career.

ISMB/ECCB 2015 will also include special sessions, poster presentations in the exhibit hall, the Student Council Symposium, and social activities. This year's conference promises to be an eagerly anticipated reboot of ISMB/ECCB with great science and collaboration at its core.

GIW/ISCB-Asia 2014

The 25th Anniversary GIW International Conference on Genome Informatics & the 4th International Society for Computational Biology Asia Conference Plaza Heisei, Tokyo, Japan, December 15 - 17

As one of the longest running annual conferences in bioinformatics or computational biology GIW has played an important role in the development of the bioinformatics community in the Asia-Pacific region since its establishment in 1990, adoption of the English language in 1993, and adoption of international site selection in 2006. Now solidly established as an international conference, GIW revisited its birthplace, Tokyo, for 2014.

2014 was an exceptionally exciting year for as the first time for GIW was held jointly with the International Society for Computational Biology, as GIW/ISCB-Asia 2014, building on three previous years of ISCB-Asia.

This year we received 86 high-quality submissions from numerous countries covering various fields of bioinformatics. We selected 35 proceedings track presentations at the conference; most of which have been published in conference associated special issues of BMC Genomics, BMC Systems Biology, IEEE/ACM Transactions of Computational Biology and Bioinformatics, and the Journal of Bioinformatics and Computational Biology, and in addition a few selected papers will be published in Bioinformatics. We also had a nine Highlights track presentations selected from 18 submissions. Moreover we were fortunate to have six world leading researchers give keynotes at the conference: Janet Kelso, Alfonso Valencia, Thomas Lengauer, Shinya Kuroda, Masami Yokota Hirai and Limsoon Wong. Finally we had approximately 100 poster presentations (including walk ons).

The conference format was unusual in allotting a long lunch time in the first two days to facilitate networking among participants and exploit the "Odaiba" waterfront location which is a tourist destination of sorts. The mechanics of the conference went smoothly and the delegates appeared happy with the conference overall and the scientific program in particular.

FASEB Names David Rocke Treasurer-elect

The Federation of America Societies in Experimental Biology (FASEB) recently named David M. Rocke, Ph.D., as treasurer-elect 2015-2016, treasurer 2016-2018. Dr. Rocke is Distinguished Professor in the Division of Biostatistics, Department of Public Health Sciences and the Department of Biomedical Engineering at the University of California, Davis, where he has been on the faculty since 1980. Rocke is a past treasurer, elected officer, and member of the board of directors for ISCB. He has served as ISCB's alternate representative to the FASEB board of directors, as well as on the FASEB Finance Committee and Public Policy Committee.

Rocke's term as treasurer-elect will begin on July 1, followed by a two-year appointment as treasurer in 2016. This is a distinguished role within the FASEB leadership and it is especially exiting to see an ISCB member hold the position.

We wish Rocke the best of luck in his new leadership role and thank him for his dedication and support for both FASEB and ISCB!

ISCB Responds to the Notice on Sustaining Biomedical Data Repositories.

by Judith Blake,
ISCB Public Affairs and Policies Committee Chair

Through a detailed listing of comments and suggestions, ISCB responded to the Notice regarding Sustaining Biomedical Data Repositories. The ISCB is the first and only society dedicated to representing the computational biology and bioinformatics community on a global scale.

As computational biologists and bioinformaticians, the ISCB community is actively engaged in the development and use of large data resources. Represented in the response, biomedical research is increasingly a data-driven enterprise, with our ability to generate and collect data outstripping our ability to analyze, interpret, and integrate it in a sustainable manner that supports its validation, reuse, and utility for advancing knowledge and translation between disciplines.

Large data resources provide essential infrastructure that benefits research funding by all NIH Institutes. Biomedical repositories that store and integrate data and knowledge bases that provide data in context of existing knowledge are critical to the biomedical enterprise. The response to this RFI is aimed at reflecting the sustainability of data repositories and knowledge bases.

Here are comments and suggestions in regards to sustainability of large data resources and knowledge bases:

Financial Models:

- Significant numbers of data resources provide essential infrastructure to all of NIH as well as other governmental agencies like FDA and EPA and NIEHS. Within NIH, support for large data repositories could be usefully managed as a trans-NIH program. Currently, individual ICs may not see sufficient benefit in investing in a given data repository when many of the benefits that are provided support the missions and investigators of other ICs.
- Funding could be a combination of intra- and extramural funding, with U41 and related grant mechanisms persisting for extramural programs. NLM might take on a more global role in oversight and management of bioinformatics resources. This might modify the resistance of Institutes to fund databases and repositories since these resources are used by researchers funded across NIH spectrum.

- It will be extremely difficult to shift to new financial models because of the extensive existing landscape of free data repositories. Databases shifting to a non-free model will likely have many free competitors, so it will be very difficult for them to compete on an equal footing. For example, NCBI provides a huge inventory of free resources; unless NCBI changes to a non-free model, it will be very difficult for non-free resources to compete with NCBI.
- We do not support institutional/academic contracts to provide access to these repositories particularly as this model gives unlimited access to scientists within these privileged networks while marginalizing and restricting access by scientists working at small or poorly funded research institutions around the world.

Innovation:

- Innovation and adoption of new technologies need to be supported as they emerge. The primary focus of integrative data resources should be accessibility and utility; both with respect to the upstream and downstream processes of data collection, and refinement and connectivity-based enhancement—all with goal of enabling downstream analyses and knowledge and hypothesis-generating tasks.

Evaluation:

- Common data usage and impact standards could be developed for application across all resources.
- Resource specific metrics might be developed and implemented for specific subtypes of resources allowing for tracking and evaluation of impact of these resources.
- It is critical that new review criteria be developed for data repositories and knowledgebases. It is highly problematic to apply the same review criteria to hypothesis-driven research as to a data repository, for example, depending on the maturity of the repository and the rate of change in the data it captures, the need for innovation may be relatively low.
- New review and evaluation mechanisms must be developed that clearly separate the review of the core mission and need for a data repository, from the review of specific new initiatives proposed by a data

repository. If the core mission and need are sound, funding of core operational aspects of the data repository should continue without interruption, because interruption could slow or halt the research of thousands of scientist users. Current review procedures allow relatively small blemishes in an otherwise sound project to derail the entire project.

Best Practices:

- Use of shared metadata and data standards is essential.
- Improvement in publication practice to include digital tagging for taxa studied, assay types, ORCHID IDs of authors, and other data standards would promote more robust and efficient integration of new data into the repositories.
- Data sharing mechanisms and procedures need to be clarified for investigators.

Partnerships:

- Partnerships with private and public sectors are essential in order not to reinvent the wheel in dealing with Big Data. However, public funding should encourage the use and the development of open source software that can be freely developed and shared across communities.

Technical:

- Some metadata standards could be required for all participating resources. These could include: Persistent identifiers of a limited set for different data types; and alignment and cross-referencing of biomedical ontologies.
- Common core repository for ubiquitously shared elements such as proteoforms for human and major model organisms would maximize access and utility of these primary components of most research projects.

Human Capital:

- Large data resources, as any repository, will go stale if not updated in terms both of semantic types and of data formats and structures. This will required appropriate staff to maintain and update any repositories and knowledgebases.
- Even automated and semi-automated curation and data representations requires skilled bioinformaticians and data analysts, typically Ph.D. level biologists with basic programming/scripting skills, to develop rule-based automated data load systems and to curate quality control aspects of the data.
- Some savings can be realized with improvements in the biomedical publishing industry to require metadata tagging prior to publications such as basic identification of experimental organism and of molecular elements such as proteins or RNAs that are under investigation.

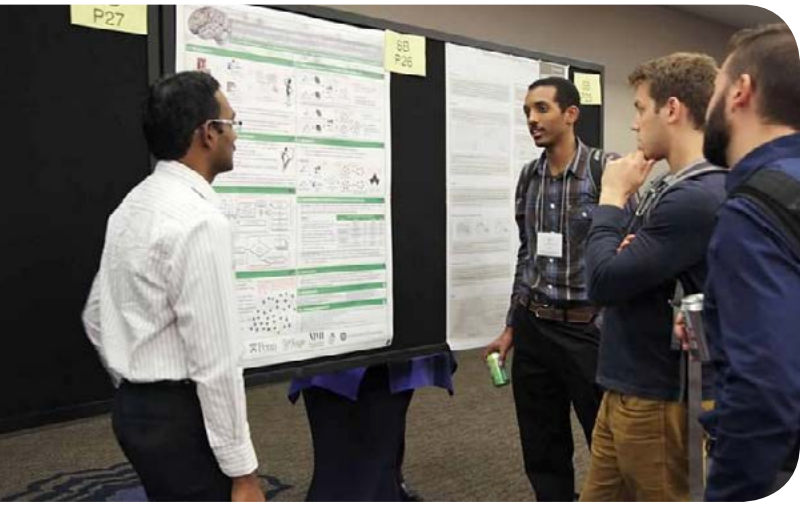
Life Cycle:

- Large data resources, with dedicated staff, will evolve in concert with emerging technologies, formats, and research interests. Leadership for these resources must recognize and support the ability to add and drop components of the resource in response to changes in the research landscape and funding support.
- Thus, there is no defacto necessity for any given resource to be retired. However, legacy data may be archived. This archiving, again, is only useful if dedicated staff maintain and update formats and continue to store the data in retrievable formats with current metadata standards applied.

Report on the 2014 RECOMB/ISCB Conference on Regulatory and Systems Genomics, with DREAM Challenges & Cytoscape Workshops

by Christopher M. Williams

When the human genome was first published in 2003, hopes were high that it would accelerate the identification of disease-causing mutations and more effective therapies. But although there has been a remarkable expansion in the discovery of genomic alterations that correlate with increased disease risk, translating these findings into new treatments has been slow. In part, this is because genome sequence is just one input into a complex system that regulates gene expression and protein activity. Noncoding DNA and elements such as transcription factors, RNA binding proteins, protein shape, and epigenomic alterations — to name a few — all play important roles in orchestrating how cells function. As scientists have gained new technologies for studying such phenomena on a genome-wide scale, there has been a growing sense that delivering on the promise of genomics will only be possible once a more comprehensive, systems-wide picture is available of how such regulatory networks function.



On November 9-14, 2014, scientists working toward this goal gathered in San Diego for the 7th Annual RECOMB/ISCB Conference on Regulatory and Systems Genomics, with DREAM Challenges & Cytoscape Workshops. The meeting highlighted new experimental and computational techniques, recent biological discoveries they have delivered, and the challenges that face the field in the coming years. The conference's four tracks featured talks and posters on systems biology, regulatory genomics, the Dialogue for Reverse Engineering Assessment and Methods (DREAM), and Cytoscape, an open source software platform for network analysis and visualization. By bringing these diverse communities together, the conference aimed to promote interdisciplinary dialogue that will advance understanding of genome function at the systems level.

Experimental methods

Research in systems biology and regulatory genomics has become possible with the arrival of high-throughput technologies for studying biology on a genome-wide scale. These include increasingly common techniques like genome, exome, and RNA sequencing, as well as complementary tools such as CHiP-seq, CLiP-seq, and DNAse-seq that, when combined with sequence data, provide additional functional information. In addition, systems and synthetic biology have produced approaches for systematically perturbing biological systems in ways that make it possible to dissect regulatory networks. Many of the talks at the conference focused on such techniques and analytical approaches for interpreting the data they generate.

In a keynote address to the DREAM and systems biology tracks, William Hahn (Dana-Farber Cancer Institute, Harvard Medical School) pointed out that although traditional genome studies can pinpoint copy number variations associated with disease, they typically have little to offer in explaining these mutations' biological function. Complicating matters further is the fact that the majority of cancers exhibit a "long tail" of rare mutations that do not rise to the top of statistical association studies. In an effort to gain insights about such mutations, the Hahn Lab is conducting genome-wide functional assays in search of essential genes and synthetic lethals, using RNA interference, CRISPR/Cas9 genome engineering, and proteomics approaches. In each case they begin with a list of genes identified for their associations with cancer, perturb a series of cell lines, and track the effects on gene expression or protein-protein interactions. Experiments so far have both pinpointed known cancer-driving alleles and previously unknown alleles that were subsequently shown to lead to disease, and have revealed the relative strengths and weaknesses of the RNAi and CRISPR approaches. The lesson, Hahn suggested, is that using a combination of experimental approaches offers "a great starting point for thinking about which genes to pursue and how to place them in biology."

Jay Shendure (U. Washington) is also interested in developing new technologies for interpreting the effects of genomic variation. Using a variety of massively parallel functional assays, his lab is prioritizing genome-wide coding and noncoding annotations to identify gene variants that are most likely to cause disease. In one assay, researchers use pools of synthetic enhancers to make every possible mutation of a native enhancer and ascertain its consequences for gene expression. In another assay, called Programmable Allelic Series, they use a panel of array-derived oligonucleotides in a parallel mutagenesis reaction to perform every amino acid swap for a protein of

interest and then observe its effects. They have also experimented with saturation editing of essential genes, making comprehensive libraries of nucleotide substitutions and watching their effects. Shendure's vision is that systematically interrogating the genome in this way could help to interpret variants of uncertain significance, and generate large datasets for training models of genome regulation.

The role of big data

In parallel with new experimental methods, systems biology and regulatory genomics have also become possible with the development of large biological data sets. How to gain biological insights from them, as well as the significant technical challenges of working with big data, comprised another key focus of this year's conference.

One valuable resource has been the Genotype-Tissue Expression Project (GTEx), an NIH Common Fund initiative that includes whole genome sequence, RNA-seq, and other kinds of data across 50 tissue types. Quantitative human geneticist Nancy Cox (Vanderbilt U.) is using computational approaches to analyze eQTL data from GTEx and other resources for genome variation that causes gene expression changes associated with disease phenotypes. Her lab's approach has identified several examples of tissue-specific genetic variation that leads to expression variation, including in type 2 diabetes and rheumatoid arthritis. Recently, Cox moved from the University of Chicago to Vanderbilt University, whose BioVU biobank contains nearly 200,000 patient DNA samples and 30 years of medical records. By combining this wealth of clinical data with her lab's methods for analyzing GTEx, Cox anticipates new opportunities for identifying mechanistic drivers of disease.

Another widely used genomics data resource is the Cancer Genome Atlas (TCGA). Ilya Shmulevich (Institute for Systems Biology) has been one of the leaders in its development, taking what he calls a "no technology left behind" approach to collecting data on primary tumor samples. The result is a comprehensive feature matrix covering multiple tumor samples across approximately 30 tumor types that includes data on DNA sequence, gene expression, DNA methylation, microRNA expression, copy number alterations, somatic mutations, clinical features, and other information. Using software called Regulome Explorer, his group has also identified associations within the data that point to signatures of cancer subtypes. At the conference he discussed projects in which they have used Regulome Explorer to stratify papillary thyroid carcinoma and gastric cancer into subtypes based on their molecular features.

Although new technologies and resources such as GTEx and TCGA offer many new opportunities, managing and analyzing big data presents a host of logistical problems. **As Lincoln Stein**

(Ontario Institute for Cancer Research) explained, the recent proliferation of inexpensive, next-generation experiments has far outstripped researchers' ability to use the data they generate. A lack of standards for handling metadata complicates the integration of data sets from different sources, while computational workflows for analyzing data often fail or are difficult to port between different computing environments. Furthermore, wildly disparate legal requirements for using biological data create a legal minefield that computational biologists navigate at their peril. Stein presented a set of ideas for addressing these challenges, suggesting that future genomics databases should be structured as a series of cloud computing environments, each of which contains both data and algorithms for analyzing them, saved in containers with all of the necessary dependencies, libraries, and operating systems. He also voiced support for the development of streamlined data access policies, including an internationally recognized consent form for patients, as well as a certification for investigators to ensure that they adhere to data ethics policies.

Echoing some of Stein's recommendations, Shmulevich also introduced a new NCI-supported project aimed at bringing TCGA to the cloud. His hope is that this effort will improve the speed of cancer research by democratizing access to data and computing power, and facilitating better collaboration.

Regulatory models

In addition to enabling association studies, genome-wide experiments and large data sets like GTEx and TCGA have made it possible to develop accurate computational models for predicting regulatory activity. Many of the talks and posters at the conference presented new methods for dissecting various layers of regulatory networks, integrating heterogeneous data sets, using networks to identify potential biomarkers and drug targets, and accounting for dynamics in regulatory networks.

In a keynote address to the DREAM track, Andrea Califano (Columbia U.) described his laboratory's attempts to move beyond one-gene-at-a-time studies of cancer by developing

genome-wide models of gene regulation. His strategy involves searching for master regulators of cancer — bottlenecks within the regulatory networks that define specific tumor subtypes. This approach provides an alternative paradigm to oncogene addiction, he explained, as master regulators are often not themselves mutated, but integrate upstream signals to become essential drivers of cancerous phenotypes. By identifying master regulators, Califano hopes to narrow the number of genes that could possibly initiate cancer, provide mechanistic information about exactly how mutations drive cancer, and identify synergies between multiple genes that are necessary for certain cancer phenotypes. Such information, as demonstrated by Califano and other speakers at the conference, could also help to identify combination therapies and predict drug responses. Recently, the Califano Lab has begun a project to reclassify tumors in TCGA on a pan-cancer basis, based on their master regulators. He has also started N-of-1 clinical trials in which patients' tumors undergo master regulator analysis, offering a potential new approach for precision medicine.

Brendan Frey (U. Toronto) uses regulatory models to score the impact of all mutations that affect regulation. His approach involves taking DNA and RNA-seq data, and then using machine learning methods to predict outcomes such as gene expression, genome stability, and RNA splicing. Mutated genomic code is then run through a similar pipeline, with the assumption that mutations that change regulatory networks should produce a change in gene expression. In his talk he described his lab's efforts to compare the effects of different types of mutations on RNA splicing. They found that mutations in introns have much more widespread regulatory effects than had been previously appreciated, and that intronic disease mutations more than 30 nucleotides from a splice site alter splicing much more often than common variants. Because RNA splicing plays an important role in gene expression and its disruption has been connected to diseases such as spinal muscular atrophy, Frey suggested that this approach could offer a new way of understanding disease.

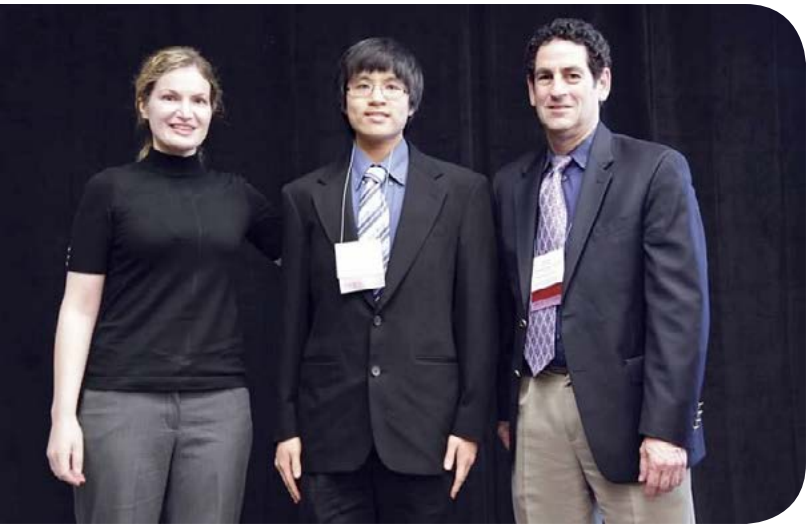
Gene Yeo (UCSD) studies RNA binding proteins (RBPs), which control the fate of messenger RNAs during RNA processing. His lab has been working to map their binding sites across the genome, and to study how RBP variants (as opposed to transcript or genome variants) change RBP regulatory networks. In recent work his lab looked at these basic biological problems in the context of Lou Gehrig's Disease (ALS), because most mutations correlated with the disease appear in genes for RBPs. He discussed their investigation of hnRBP A2/B1, which used CLiP-seq, shRNA knockdown assays, and tethering experiments. These efforts identified variants in A2/B1 that cause defects in the life cycle of RNA processing, and provided some clarity about the role of RBPs in stress granule formation, a hallmark of neurodegeneration. Yeo remarked that his approach has similar goals to those of efforts to map transcription factor binding sites, but pointed out that study of RNA binding proteins is complicated by the fact that they are found not only in the nucleus, but also throughout the cell.

In a talk that provoked an enthusiastic if chagrined response from members of the audience, Ellen Rothenberg (Caltech) discussed her lab's elegant work to understand how hematopoietic stem cells differentiate into T cells. Their findings suggest that the T cell program is more than an ordered activation of multiple genes by a master regulator, but is "maximally sloppy," offering many different combinations of regulatory options. She hypothesized that this multipotentiality exists for evolutionary reasons, as it offers functional plasticity that addresses the body's need to adapt the hematopoietic system for multiple roles. In the question and answer period that followed her talk, several attendees commented that Rothenberg's experimentally derived findings highlight current limitations in computational biology's ability to model such a dynamic system. But Rothenberg was optimistic that new high-throughput technologies for perturbing systems and single-cell techniques offer opportunities that weren't available even a few years ago.

Variability in single cells

Until recently, computational biology had to make do with data based on average measurements taken across large populations of cells. This presents a problem for understanding complex diseases like cancer, however, because bulk samples can contain widespread genomic and molecular heterogeneity among small subpopulations of cells. As William Hahn pointed out early in the conference, it is often these rare cells that enable cancer to persist and become resistant to treatment. Several keynote talks focused on methods for parsing this variability.

One recent innovation in single-cell biology has been the development mass cytometry, which enables researchers



to simultaneously tag up to 45 features in individual cells with isotope labels. The cells are fed one at a time through a narrow channel, where they are then ionized and their isotope signatures are read quantitatively. The technology's inventor, Garry Nolan (Stanford U.), gave an overview of new opportunities it offers, as well as of a new technology called multiplexed ion beam imaging (MIBI). He anticipates that MIBI should soon make it possible to build an 8 billion voxel image of a single cell's structure, including DNA and proteins. Having this degree of resolution, he suggested, will make it possible to compare clusters of cells across animals or species and provide reference maps that could be used to locate other cells. Single-cell techniques also have clinical applications, he explained, as his lab has begun creating an immune system reference map that could help predict how quickly individual patients will recover following surgery.

At the same time, a technology capable of tracking 45 parameters over millions of single cells presents enormous computational challenges. To address this problem, Nolan has been collaborating with Dana Pe'er (Columbia U.), who delivered another keynote address focusing on her lab's efforts to dissect heterogeneity in cancer. Recently, they developed an algorithm called viSNE, which reduces this 45 dimensional space into an interpretable 2-dimensional graph. By scoring each cell and then ordering the entire collection based on the amount of similarity between cells, it becomes possible to cluster them in an orderly fashion. Pe'er also argued that single cell approaches enabled by technologies such as mass cytometry offer an advantage over genomic approaches because they measure actual cellular activity, and not just gene expression. Elsewhere in her talk, Pe'er discussed another algorithm her lab developed, called Helios, which integrates results of functional RNAi screens and other data types to predict copy number variants that drive cancer.

Amos Tanay (Weizmann Institute of Science) is using single-cell strategies to explore how information coded in the genomic sequence computes gene regulation, and proposed three key issues that make this difficult. One is that gene expression is not deterministic, but is quantitative, complex, and stochastic. To account for this property, he uses single-cell RNA-seq to sample populations of cells in an unbiased fashion. He then uses network models to organize them into cell subtypes, treating gene expression as distributions to account for sampling variances and the fuzziness of boundaries between cell types. A second problem is that cells use methylation as a memory system, reflecting different degrees of commitment along lineages of differentiation. Using a computational approach called UMI-bis to study methylation distribution in embryonic stem cells, he found that methylation is not a single process, but one that balances the addition and removal of methylation using specific enzymes. Finally, he pointed out, the genomic code is not pure information, but exists in physical space within the nucleus. Using Hi-C to catalog groups of interactions that result from protein folding, his goal is to identify "contactomeres" that reflect how accessibility affects gene expression on a full-chromosome scale.

Physical chemist James Heath (Caltech) uses single-cell proteomic data and principles from thermodynamics to develop predictive models of disease progression in cancer. His approach involves considering cells as occupying steady states that change when subjected to perturbations or changes in regulatory architecture. By compiling changes in protein expression and different measurements of single cell signaling networks into a histogram, he explained, his lab captures fluctuations of a system that indicate movement of cells from a normal steady state to a drug-resistant steady state. Unlike network-based approaches his strategy doesn't require knowing mechanistic details about the network, but can generate biologically relevant insights into the causes of state changes. He discussed applications of his approach to studying glioblastoma and melanoma, including some clinically relevant findings that the lab produced.

DREAM Challenges

As accuracy and reproducibility of quantitative approaches continue to be a concern of the computational biology community, the DREAM Challenges provide an objective way of comparing various methods. Each year, they invite computational biologists to apply their chosen methodologies to solve key problems facing the fields of systems biology and translational medicine. In a spirit of friendly competition, their solutions are assessed in order to identify the most effective strategies. This year's activities included the DREAM Rheumatoid Arthritis Responder Challenge, the ICGC-TCGA DREAM Somatic Mutation Calling Challenge, the DREAM

Acute Myeloid Leukemia Outcome Prediction Challenge, and the Broad-DREAM Gene Essentiality Prediction Challenge. At the conference, the results for each challenge were presented and best performers gave short talks discussing their approaches.

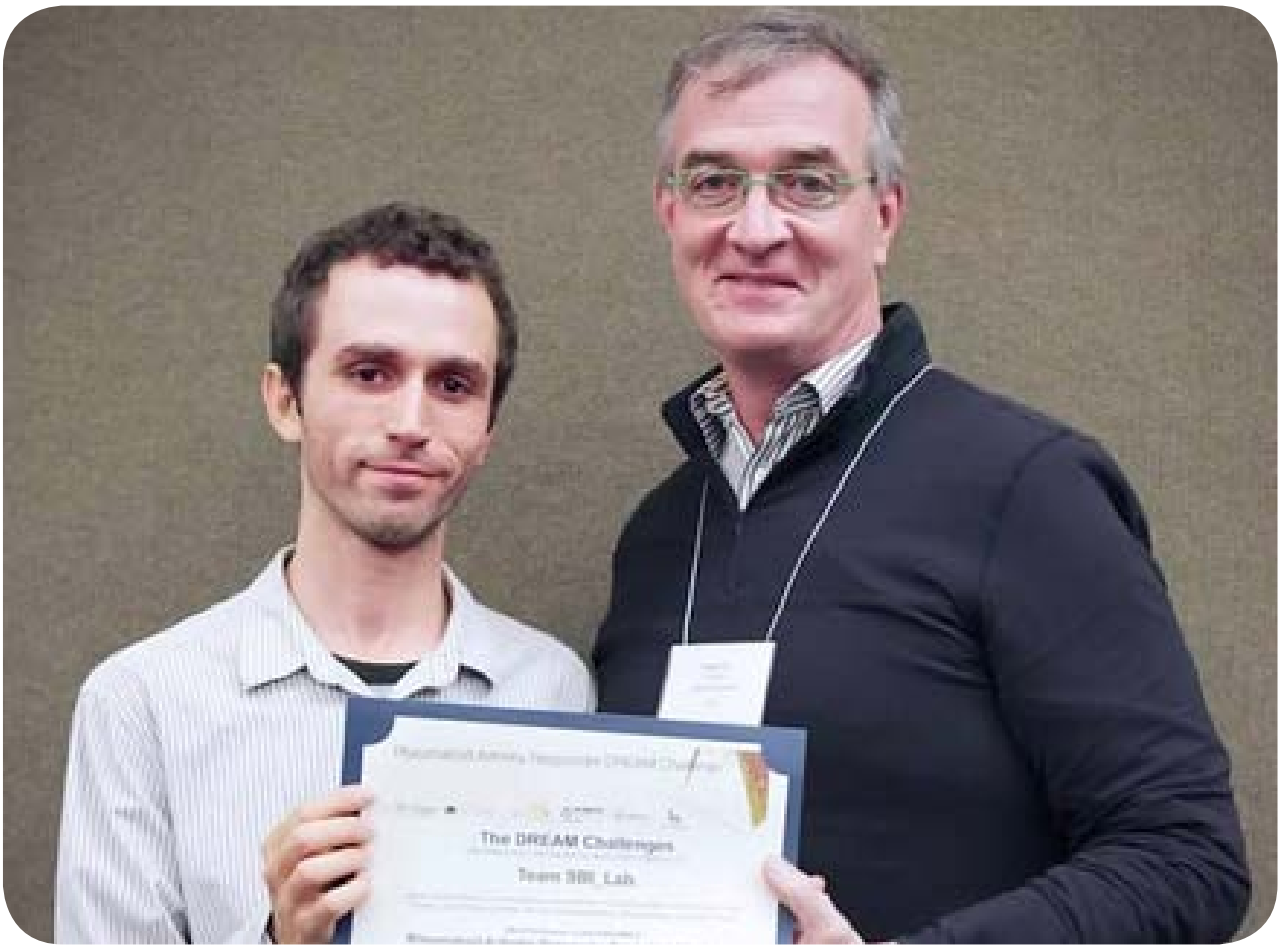
Summarizing the results, Gustavo Stolovitzky (IBM) highlighted some broader insights that emerged from the four DREAM challenges. For the AML and rheumatoid arthritis challenges, he pointed out, it was ultimately clinical data and not genomic data that proved most valuable for the best performers, raising questions about the value of genomic data for studying these diseases. For the gene essentiality challenge, he noted that gene expression proved to be the most valuable data type, instead of SNP and CNV mutations. He also highlighted the fact that the somatic mutation calling challenge provided the first opportunity for the computational community to develop algorithms focused on this problem, and should prove valuable for benchmarking progress in this nascent field.

In the end, the conference as a whole offered no easy answers to the question of how to best understand cellular activity on a systems-wide level. Nevertheless, it showcased many exciting new developments that are bringing the community closer to this goal, making it clear that as new technologies, data sets, and algorithms continue to evolve, they are producing an unprecedentedly detailed picture of how genetic systems are structured and regulated.

To download the conference program and view video from the conference, visit http://www.iscb.org/recomb-regsysgen2014.

The 8th Annual RECOMB/ISCB Conference on Regulatory Genomics and Systems Biology, with DREAM Challenges takes place on November 16-18, 2015 in Philadelphia, PA, USA. It is currently accepting submissions of papers and abstracts. To learn more, visit http://www.iscb.org.

A video about the DREAM Challenges is available at http://bit.ly/18haNhb. To learn more about this year's challenges, visit dreamchallenges.org.



The International Society for Computational Biology (ISCB) announces new ISCB Ebola Award

The Ebola virus (EBOV), causing the Ebola virus disease (EVD) maintains a high fatality that has killed over 5,000 individuals in the current outbreak from February to Nov 26, 2014. First reports assessing spreading risk have been published (Gomes, et al., 2014). Recently, 99 Ebola virus genomes from 78 patients have been analyzed (Gire, et al., 2014). To facilitate global research, the respective author team made all their data freely available. This laudable decision aligns with the ISCB Open Access policy (Lathrop, et al., 2011; Lathrop, et al., 2011).

In response to the immediate need for solutions in the field of computational biology against Ebola, The International Society for Computational Biology (ISCB) announces the *ISCB Ebola Award*. ISCB will give out the *ISCB Ebola Award*, along with a prize of $2000, at its July 2016 annual meeting (ISCB ISMB 2016, Orlando, Florida). All computational approaches should include a significant component of Ebola research. In the development of any modern drug, computational biology is positioned to contribute through comparative analysis of the genome sequences of Ebola strains, and 3-D protein modeling. Other computational approaches to Ebola include large-scale docking studies of Ebola proteins with human proteins and with small-molecule libraries, computational modeling of the spread of the virus, computational mining of the Ebola literature, and creation of a curated Ebola database. Taken together, such computational efforts could significantly accelerate traditional scientific approaches.

ISCB is dedicated to advancing the understanding of living systems through computation. ISCB now represents more than 3000 computational biologists working in over 70 countries. It organizes more than seven annual international meetings, and confers several major prizes, including the *ISCB Senior Scientist Award*, the *ISCB Overton Prize*, and the *ISCB Outstanding Contributions Award*. With the *ISCB Ebola Award* the society offers for the first time an award for a specific scientific objective thereby acknowledging the urgency of action to fight a rising challenge.

Gire, S.K., et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science 2014;345(6202):1369-1372.

Gomes, M.F.C., et al. Assessing the International Spreading Risk Associated with the 2014 West African Ebola Outbreak. PLOS Currents Outbreaks 2014.

Lathrop, R.H., et al. ISCB public policy statement on open access to scientific and technical research literature. Bioinformatics 2011;27(3):291-294.

Lathrop, R.H., et al. ISCB Public Policy Statement on Open Access to Scientific and Technical Research Literature. PLoS computational biology 2011;7(2):e1002014.

Recruit - Job Search - Network

ISMB/ECCB 2015 CAREER FAIR

http://www.iscb.org/ismbeccb2015-career-fair

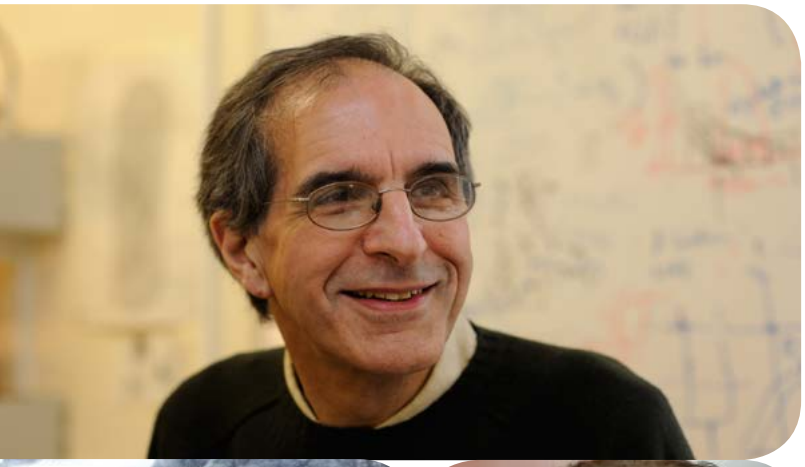
Tweet Alert – ISCB moves all Twitter handles under @ISCB

ISCB is working to streamline our Twitter communication and hash tags. Instead of having individual accounts for each major event, ISCB is going to use one main twitter handle - @ISCB. Each event will have it's own hash tag like #ISMBECCB15, #GLBIO, #ISCBNGS, #Rocky. This will allow us to track the conversation and keep you better informed.

Follow us today!

Cyrus Chothia, Curtis Huttenhower, and Larry Hunter Named 2015 ISCB Award Winners

by Christiana N. Fogg, Ph.D.



The International Society for Computational Biology (ISCB) is pleased to announce the winners of the 2015 Senior Scientist Accomplishment Award, Overton Prize, and the inaugural Outstanding Contributions to ISCB Award. Cyrus Chothia, Emeritus Group Leader at the Medical Research Council (MRC) Laboratory of Molecular Biology in Cambridge, England has been selected as the winner of the 2015 Senior Scientist Accomplishment Award. This year's Overton Prize honors Curtis Huttenhower from the Harvard School of Public Health, and Larry Hunter from the University of Colorado School of Medicine is the first recipient of Outstanding Contributions to ISCB Award.

The ISCB Senior Scientist Accomplishment Award recognizes leaders in computational biology and bioinformatics for their significant contributions to these fields through research, education, and service. Cyrus Chothia was selected as the 2015 recipient for his groundbreaking work using computation to understand protein structure and function and the evolution of genomes.

Chothia is well known for using computation to study protein structure, and his early work showed that relatively simple principles govern the structure of proteins, regardless of the structural complexity. His research has been critical to understanding and classifying proteins based on structural folds, and he has shown that changes to a protein sequence can be accommodated by structural shifts. More recently, Chothia developed computational approaches based on his knowledge of protein structure to understand how gene duplication and recombination between particular domains drives genome evolution. Chothia's illustrious career includes election as a Fellow of the Royal Society in 2000. He has mentored numerous students and postdoctoral fellows, and many are now rising leaders in their respective fields. Chothia's work throughout his career has been instrumental to the birth of the fields of structural bioinformatics and computational genomics.

The Overton Prize recognizes early or mid-career scientists who are emerging leaders in computational biology and bioinformatics for their accomplishments in research, education, and service. The Overton Prize was instituted in 2001 to honor the untimely loss of G. Christian Overton, a leading bioinformatics researcherand founding member of the ISCB Board of Directors. Curtis Huttenhower is this year's winner of the Overton Prize for his groundbreaking research on microbial communities, with a focus on the human microbiome.

Huttenhower is an Associate Professor of Computational Biology and Bioinformatics at the Harvard School of Public Health. He has worked on developing novel computational tools to analyze the large, complex datasets associated with microbial communities and NIH Human Microbiome Project. Huttenhower's research has provided new insights into how microbial communities impact human health and disease. His research potential has been recognized through the receipt of the Presidential Early Career Award for Scientists and Engineers and an NSF CAREER Award.

2015 marks the inaugural recognition of an ISCB member with the Outstanding Contributions to ISCB Award for his or her outstanding service contributions toward the betterment of ISCB through exemplary leadership, education, and service. Larry Hunter is the first winner of this award for his instrumental role in the foundation of ISCB as a scientific society.

Hunter is a Professor and Director of the Center for Computational Biology at the University of Colorado School of Medicine. His computational biology research interests include biomedical text mining and knowledge-based computational techniques for analysis of high-throughput data. Hunter began his career

as a programmer at the U.S. National Library of Medicine (NLM), where he developed a database of researchers interested in artificial intelligence and molecular biology. He invited researchers from this database to a joint NLM-NSF meeting on artificial intelligence in molecular biology in 1992. This meeting developed into the Intelligent Systems in Molecular Biology Meeting (ISMB). By 1996, ISMB had emerged as the premier meeting for computational biology research, and members of previous ISMB steering committees concluded that this unique interdisciplinary field needed its own professional society. This group of committee members created the International Society of Computational Biology and elected Larry Hunter as its first president. Hunter has gone on to serve ISCB in many other capacities and continues to be closely involved with ISCB. ISCB will present the Senior Scientist Award, Overton Prize, and Outstanding Contributions to ISCB at the joint 23rd Annual ISMB/14th Annual European Conference on Computational Biology (ECCB) being held in Dublin, Ireland, July 10-14th, 2015. Chothia, Huttenhower, and Hunter will present keynote talks during ISMB/ECCB 2015.

Full bibliographical articles profiling the award recipients will be available in the ISMB/ECCB 2015 focus issue of the ISCB newsletter later this year, as well as the ISCB Society Pages in PLOS Computational Biology and OUP Bioinformatics.

Tenth Annual Great Lakes Bioinformatics Conference: GLBIO 2015

An official conference of the International Society for Computational Biology. Organized by the Great Lakes Bioinformatics Consortium. Co-hosted by Purdue University and the International Society for Computational Biology

The Great Lakes Bioinformatics Conference:
~ Provides an interdisciplinary forum for the discussion of research findings and methods.
~ Fosters long-term collaborative relationships and networking opportunities within the domain of computational approaches to biology.

GLBIO has established a strong reputation for building relationships among a nationally prominent bioscience research community, showcasing the North American Great Lakes region as a perfect place to conduct computer -aided research.

The ISCB's Tenth Great Lakes Bioinformatics Conference will be held May 18-20 at Purdue University. The event will offer outstanding keynote speakers, tutorials, workshops, oral presentations and posters.

Featured Keynote Speakers:
Nancy Cox - Large-scale Data Integration: Genomes, Transcriptomes, and Electronic Medical Records
Mona Singh - Data driven approaches for uncovering variation in protein interaction networks
Tandy Warnow - Grand Challenges in Phylogenomics
Jian-Kang Zhu - Regulation of DNA methylation in plants

View the conference agenda:
http://www.iscb.org/agenda-program/glbio2015-detailed-agenda

KEYNOTE SPEAKERS

Eileen Furlong
European Molecular Biology Laboratory, Heidelberg, Germany

Michael Levitt
2013 Nobel Laureate – Chemistry; Stanford University, United States

Kenneth H Wolfe
University College Dublin, Ireland

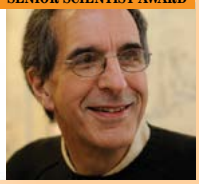
2015 OVERTON PRIZE AWARD

Curtis Huttenhower
Harvard University, United States

2015 ISCB FELLOWS KEYNOTE

Amos Bairoch
University of Geneva; Swiss Institute of Bioinformatics, Switzerland

2015 ISCB ACCOMPLISHMENT BY A SENIOR SCIENTIST AWARD



Cyrus Chothia
MRC Laboratory of Molecular Biology, Cambridge, United Kingdom

THE SCIENCE

With more than 200 talks by renowned researchers and over 800 posters on computational biology and bioinformatics, ISMB/ECCB 2015 is scientifically the top meeting of 2015!

INFORMATION & TRAINING

ISMB/ECCB 2015, a thematically organized conference, is your premier source for information and taking. ISMB/ECCB 2015 features pre-conference Special Interest Groups (SIGs) and Satellite Meetings, with additional Applied Knowledge Education Sessions (AKES).

CONNECT & COLLABORATE

Whether scanning the jobs board for a career opportunity, interacting with one of our exhibitors and sponsors, or discussing research at a poster session, ISMB/ECCB 2015 provides you a global networking atmosphere with more than 1500 delegates from 50+ countries.

REGISTRATION

March 24, 2015	Registration Opens
June 6, 2015	Early Registration Cut-off Date
June 27, 2015	Online Registration Closes

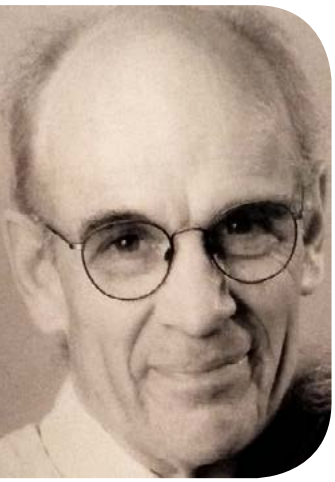
WWW.ISCB.ORG/ISMBECCB2015

ISCB INTERNATIONAL SOCIETY FOR COMPUTATIONAL BIOLOGY
OFFICIAL CONFERENCE

Peter H. Sellers, among the earliest researchers on DNA and protein sequence comparison, died of cancer on November 15, 2014, at the age of 84.

An obituary was published in The Philadelphia Inquirer [1]. Here, I offer a brief, personal perspective.

by Stephen Altschul



I first met Peter Sellers over thirty years ago, on May 28, 1984 to be precise. I was at the time a graduate student in applied mathematics at MIT, and was interested in finding a dissertation project in computational biology. The first comprehensive text book on sequence comparison [2] had just been published, from which I had gleaned the elements of this new field and the names of a number of its pioneers, including Peter. I learned that several would be speaking at a symposium planned for that spring's AAAS meeting in New York, which I eagerly made plans to attend. The symposium [3] had been organized and was being co-chaired by David Lipman from the NIH, who recently had published with John Wilbur a fast DNA and protein database search program [4]. Among the meeting's six speakers, I found Peter the easiest to approach, and he invited me to come by his office at Rockefeller University the following day.

A mathematician by training, Peter had first become interested in sequence comparison about a decade earlier, after attending a talk by the physicist Stan Ulam and solving a problem he had posed. Recently, Peter had struck up a collaboration with Bruce Erickson [5], a protein chemist interested in the synthesis of artificial peptides, who was a fellow researcher at Rockefeller University. My meeting with Peter was pleasant, but it did not shake loose any ideas for research projects. It was timed, however, so we could attend a lecture by Mike Zuker on the prediction of RNA folding, a reprise of his AAAS talk of the previous day, and we proceeded afterwards to the Rockefeller faculty club with a small group for lunch. There, I was seated next to Bruce Erickson, with whom I struck up a lively conversation. After a visit to his lab following lunch, he invited me to come work with him over the summer, a chance that I jumped at. During those months I was able to solve a problem he had been thinking about for a while, resulting in my first published paper [6] and an invitation to join Bruce's lab more permanently as a research assistant. I was able to arrange this with my thesis advisor, and spent most of the next three years with Bruce at Rockefeller, but frequently consulting with and inspired by Peter.

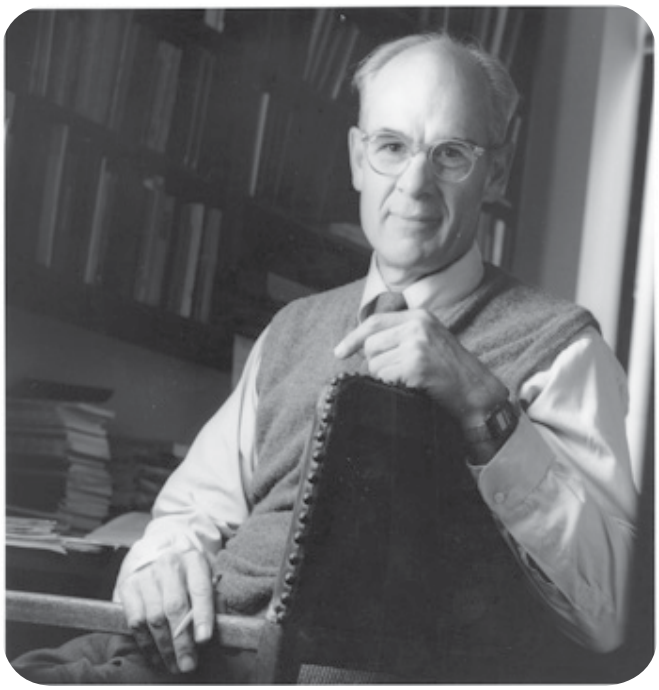
Peter Sellers' early work in sequence comparison is probably best encapsulated in three papers that take a mathematically rigorous approach to sequence alignment. In the first [7], he defines a distance measure between sequences, which he proves to be a metric, and he also describes and proves valid an algorithm for calculating it. This is, basically, what is today commonly known as the Needleman-Wunsch algorithm, although a careful reading of Needleman & Wunsch's paper [8] shows that they in fact describe a different algorithm. It is really Sellers' algorithm [7], or a related one by Sankoff [9], that is currently used. Peter's second alignment paper [10], with Bruce Erickson, describes semi-global alignment, which finds approximate occurrences of a short sequence within a longer one. His third paper [11] describes what can be understood as a more general version of the Smith-Waterman algorithm [12].

In brief, after the definition of optimal global alignment in the early '70s, it took many years for practical definitions and algorithms for local sequence alignment to be formulated [11-13]. Once this was achieved, local alignment served as the foundation for the development of rapid and popular heuristic protein and DNA database search programs [3,14,15]. Peter took the approach of defining the concept of "local optimality", and then described an algorithm for rigorously finding all and only the locally-optimal local alignments of two sequences. When his algorithm is restricted to finding the single best such alignment, it can be shown to be equivalent to that of Smith & Waterman [12]. The complete version of Peter's algorithm was conjectured to run in time O(mn), and I was able to develop a modification [16] that could be proved to do so.

Throughout my time at Rockefeller, Peter Sellers was a guiding light for my work with Bruce Erickson. Peter and Bruce made an unusual team. Bruce had the entrepreneurial spirit required of a modern lab head, with a constant need to write grants and raise funds. Peter was a much gentler soul, perhaps heavily influenced by the Philadelphia Quaker tradition. He was happy to work on his own, pursuing mathematical rigor, and was not focused exclusively on practical results. He was never a self-promoter, and his contribution to the field of sequence comparison is unfortunately frequently overlooked. As Peter's obituary [1] states, he had many interests beyond

mathematics and science. He was descended from the Peale family of 18th and 19th century American artistic fame, and his wife and several of his children have strong artistic interests. Peter served for some time teaching in Africa. He and his family removed each summer to their rustic second home on Mount Desert Island, where I had the pleasure of visiting them several times. There, having spent many years hand-building a fine wooden sloop in a Pennsylvania barn, he enjoyed many more sailing it with his wife around the coast and islands of Maine. He was happy to be able to return to Maine this past summer, although his flagging strength had put an end to his sailing days. It was a privilege to have known Peter for the past thirty years. He was a true gentleman and a true friend.

[1] Cook, B.L. (2014) "Peter H. Sellers, 84, scientist." http://articles.philly.com/2014-11-25/news/56459790
[2] Sankoff, D. & Kruskal, J.B. (eds.) (1983) Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison. Addison-Wesley, Reading, MA.
[3] Miura, R.M. (ed.) (1986) Some Mathematical Questions in Biology - DNA Sequence Analysis. American Mathematical Society, Providence, RI.
[4] Wilbur, W.J. & Lipman, D.J. (1983) "Rapid similarity searches of nucleic acid and protein data banks." Proc. Natl. Acad. Sci. USA 80:726-730.
[5] Felix, A.M (2000) "Bruce W. Erickson: In memoriam." In Peptides for the New Millennium, G.B Fields, J.P. Tam & G. Barany (eds.), p. 19, Kluwer Academic Publishers, New York, NY.
[6] Altschul, S.F. & Erickson, B.W. (1985) "Significance of nucleotide sequence alignments: a method for random sequence permutation that preserves dinucleotide and codon usage." Mol. Biol. Evol. 2:526-538.
[7] Sellers, P.H. (1974) "On the theory and computation of evolutionary distances." SIAM J. Appl. Math. 26:787-793.
[8] Needleman, S.B. & Wunsch, C.D. (1970) "A general method applicable to the search for similarities in the amino acid sequences of two proteins." J. Mol. Biol. 48:443-453.
[9] Sankoff, D. (1972) "Matching sequences under deletion-insertion constraints." Proc. Natl. Acad. Sci. USA 69:4-6.
[10] Erickson, B.W. & Sellers, P.H. (1983) "Recognition of patterns in genetic sequences." In Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison, D. Sankoff & J.B. Kruskal (eds.), pp. 55-91, Addison-Wesley, Reading, MA.
[11] Sellers, P.H. (1984) "Pattern recognition in genetic sequences by mismatch density." Bull. Math. Biol. 46:501-514.
[12] Smith, T.F. & Waterman, M.S. (1981) "Identification of common molecular subsequences." J. Mol. Biol. 147:195-197.
[13] Goad, W.B. & Kanehisa, M.I. (1982) "Pattern recognition in nucleic acid sequences. I. A general method for finding local homologies and symmetries." Nucl. Acids Res. 10:247-263.
[14] Pearson, W.R. & Lipman, D.J. (1988) "Improved tools for biological sequence comparison." Proc. Natl. Acad. Sci. USA 85:2444-2448.
[15] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410.
[16] Altschul, S.F. & Erickson, B.W. (1986) "Locally optimal sub-alignments using nonlinear similarity functions." Bull. Math. Biol. 48:633-660.



Update from FASEB's Office of Public Affairs

By Yvette Seger, PhD, Director of Science Policy, FASEB

The Federation of American Societies for Experimental Biology (FASEB) continues to work on behalf of its 27 member societies to develop and promote policies to advance research and education in the biological and biomedical sciences. A key challenge continues to be predictable and sustainable federal funding for research. Following the dramatic increases in its budget from 1999 to 2003, funding for the National Institutes of Health (NIH) has failed to keep pace with rising costs. FASEB's efforts to convince Congress of the importance of investment in basic biomedical research have produced modest increases. More importantly, congressional understanding of the need to increase funding for biomedical research is on the rise, with several new bills proposing strategies to increase research funding. FASEB continues to drive discussion of the need for federal research funding through its testimony to congressional committees, visits to congressional offices, and public outreach initiatives.

While funding issues are a critical component of FASEB's message, FASEB recognizes that this is not the only challenge faced by the research community. Changes in the funding and regulatory environments have forced all components of the research community – from individual researchers to institutions and federal agencies – to adapt. Over the past two years, FASEB engaged its leadership in an extensive, data-driven examination of the major challenges facing the biological and medical research enterprise and potential opportunities to alleviate them. Through its Science Policy Committee, FASEB conducted analyses and developed recommendations that were the focus of three roundtable discussions with representatives from FASEB societies, funding agencies, research organizations, and other stakeholder groups. Sustaining Discovery in Biological and Medical Sciences was released in January 2015 and presents key themes that emerged from the analyses and discussions and offers recommendations to maximize federal funding, optimize the research workforce, and improve the mechanisms used to support research.

In 2014, FASEB issued comments on agency proposals and policies for data sharing, clinical research, grant application and progress reports and new mechanisms to fund research. Thus far in 2015, the Federation has tackled a similarly broad range of issues, submitting comments on proposed changes to requirements for data collection and results submission for clinical trials supported by funds from the U.S. Department of Health and Human Services, NIH's principles and guidelines for reporting preclinical research in scientific journals, changes to the NIH biosketch, and draft legislation that could affect several aspects of the research enterprise. Across disciplines, researchers share common goals and many of the same challenges. Therefore, in addition to Federation-level efforts, FASEB provides resources, such as federal funding recommendations to state and district-level factsheets to assist individual investigators initiate a dialog about the importance of basic biomedical research with their colleagues and elected officials.

Members of the International Society for Computational Biology (ISCB) continue to play a critical role in FASEB leadership and ensure the consideration of data science and computational resources in policy discussions. Harel Weinstein, DSc, serves as the Chair of FASEB's Data Science and Informatics Subcommittee and represents ISCB on the Science Policy Committee. Judith Blake, PhD, represents ISCB on the FASEB Board of Directors and provides input as a member of the Data Science and Informatics Subcommittee. In addition to his role as a Board Advisor and member of the Data Science and Informatics Subcommittee, David Rocke, PhD, was recently elected as FASEB's Treasurer-Elect. If you would like to learn more about opportunities to represent ISCB on a FASEB committee or subcommittee, please contact Diane Kovats for more information.

Student Council Symposiums – Why young researchers should not miss out on them

For the last decade, the Student Council Symposiums (SCSs) have become the flagship events of the ISCB Student Council. The Symposiums are great networking opportunities that give students and young researchers the opportunity to interact in an informal setting through fun ice breakers, keynote addresses, socials events, student talks, and regular poster sessions.

Latin American-SCS (LA-SCS) was held for the first time in Belo Horizonte, Brazil on October 27th, 2014. Organizing a Symposium in a new region is always a challenge. However thanks to a team of incredibly enthusiastic students, the event was a great success. Furthermore the seed for a student network at a Latin American level was planted with ongoing local efforts to create new Regional Student Groups (RSGs) in different countries.

"My travel plans to Brazil started almost by chance, and it's one of the best decisions I ever made. Having the chance to interact with speakers, LA-SCS staff and attendants was very prolific, and winning the best poster award gave me a valuable push to continue developing myself in this amazing discipline. I am very thankful to the organizing committee and all the great friends that I made at the LA-SCS in Brazil."
- Javier Caceres Molina, LA-SCS attendee

The ISCB Student Council Symposium-Africa (SCS-Africa) in its first edition, was organized in the context of the ISCB/ASBCB

Bioinformatics Conference 2015. Young students and researchers from around the African continent gathered to share their work and interests as well as to network on a student level. The symposium opening session started with a lecture from Dr. Manuel Corpas on developing next-generation computational biologist in Africa through RSGs. It also featured a training session on cloud computing for computational biology given by Yassine Souilmi and was concluded by a networking and social event.

"I attended the SCS in Tanzania and I was one of the luckiest having the chance to present my work. I found it really exciting meeting other students from all over Africa, sharing their ideas

and showing their work. It was also a great opportunity to explore other domains in the field of computational biology."
- Mohamed Alibi, SCS Africa attendee, Best Oral Presentation winner

The eleventh Student Council Symposium (SCS) will run ahead of the ISMB/ECCB 2015 conference in Dublin, Ireland. The event provides an invaluable opportunity for students to present their work to their peers as posters or oral presentations. This year, SCS will host two pioneering computational biologists Prof. Ruth Nussinov and Prof. Des Higgins, as keynote speakers. It is still not too late to participate; the call for abstracts from student and post-doctoral researchers is still open at http://symposium.iscbsc.org/.

The ISCB Student Council would like to thank all our volunteers and advisors without whom these events would not be possible. The Student Council furthermore remains grateful to the ISCB, the ASBCB and associated communities for their unwavering support. We look forward to meeting you all in Dublin or at one of our other events!

Upcoming Conferences

Great Lakes Bioinformatics Conference 2015 (GLBIO 2015)
United States - Indiana - West Lafayette
Purdue University
May 18 - 20, 2015
www.iscb.org/glbio2015
The 5th GLBIO provides an interdisciplinary forum for the discussion of research findings and methods. An important goal for the conference is to foster long term collaborative relationships and networking opportunities within the domain of computational approaches to biology.

ISMB/ECCB 2015
Ireland, Dublin
July 10 - 14, 2015
www.iscb.org/ismbeccb2015
As the world's premier conference on computational biology, ISMB/ECCB attracts top international scientists and key decision makers in the life sciences — experts in areas such as computer science, molecular biology and medicine, mathematics and statistics — from the world's largest and most prestigious research institutions.

RECOMB/ISCB Conference on Regulatory and Systems Genomics, with DREAM Challenges 2015
Philadelphia, Pennsylvania, USA
November 16 - 18, 2015
www.iscb.org/recomb-regsysgen2015
Now in its eighth year, the RECOMB/ISCB Conference on Regulatory and Systems Genomics, with DREAM Challenges is one of the premier annual meetings in the fields of regulatory genomics and systems biology.

13th Annual Rocky Mountain Bioinformatics Conference
United States - Colorado - Aspen/ Snowmass
December 10 - 12, 2015
www.iscb.org/rocky2015
13th Annual Rocky Mountain Bioinformatics Conference offers an opportunity to focus on regional development in the computational biosciences. Representing a broad spectrum of universities, industrial enterprises, government laboratories, and medical libraries from around the world, the meeting is a chance to get to know your colleagues near and far, seek collaborative opportunities, and find synergies that can drive our field forward.

Affiliated Conferences

GIW/InCoB-Asia 2015
Japan – Tokyo
Sep 09 through Sep 11, 2015
http://www.jsbi.org/giw-incob2015/
GIW/InCoB 2015 is the first joint conference between GIW and IncoB, two prominent international bioinformatics conferences in East Asia with a combined history of 40 years.
Event Registration: 2015-05-20 through 2015-09-02
ISCB Member Discount: 15%

NETTAB & Integrative Bioinformatics 2015 Joint Conference
Italy - BA - Ruvo di Puglia
Oct 14, 2015 through Oct 16, 2015
http://www.igst.it/nettab/2015/
The challenge for Integrative Bioinformatics is to capture, model, integrate and analyze large quantities of high dimensional data scattered across thousands of biological databases and hundreds of scientific journals in order to provide systematic insights into complex biological systems.
Early Registration Deadline: 2015-09-17
ISCB Member Discount: 30 EUR

Big Data in Health Care - Challenges, Innovations and Implementation
Oct 28, 2015 through Oct 29, 2015
http://bigdata2015.uni.lu
The main objective of the Symposium is to bring together worldwide experts from academia, industry, doctors, policy makers and patient organisations in the field of Big data in Health Care to exchange about state-of-the-art research and technologies.
ISCB Member Registration Discount 20%:

Pacific Symposium on Biocomputing (PSB) 2016
United States – HI – Kohala Coast
Jan 04 through Jan 08, 2016
http://psb.stanford.edu/
The Pacific Symposium on Biocomputing (PSB) 2016 is an international, multidisciplinary conference for the presentation and discussion of current research in the theory and application of computational methods in problems of biological significance.
Event Registration: 2015-08-03 through 2016-01-08
ISCB Member Discount: 50 USD

Other Conferences & Events of Interest

EMBO Conference Series: Chromatin and Epigenetics
Germany – Heidelberg
May 06, 2015 through May 10, 2015
Hosted by: European Molecular Biology Laboratory
http://www.embl.de/training/events/2015/CHR15-01/index.html

Applied Bioinformatics and Public Health Microbiology
United Kingdom – Cambridgeshire – Cambridge
Hosted by: Wellcome Trust
May 06, 2015 through May 08, 2015
http://registration.hinxton.wellcome.ac.uk/display_info.asp?id=474

2nd Symposium on Complex Biodynamics & Networks
Japan – Tsuruoka
Hosted by: Institute for Advanced Biosciences, Keio University, Japan
May 11, 2015 through May 13, 2015
http://www.c-bio.org/2015/

EMBL Conference: Biology and Pathology of the Malaria Parasite
Germany – Heidelberg
Hosted by: EMBL Heidelberg
May 11, 2015 through May 13, 2015
http://www.embl.de/training/events/2015/BMP15-01/

NIMBioS Investigative Workshop: Malaria-Leishmaniasis Co-infection
United – TN – Tennessee – Knoxville
Hosted by: National Institute for Mathematical and Biological Synthesis (NIMBioS)
May 26, 2015 through May 28, 2015
http://www.nimbios.org/workshops/WS_coinfection

Workshop on Statistical Learning of Biological Systems from Perturbations
Switzerland – Ticino – Ascona
Hosted by: Niko Beerenwinkel, Peter Bühlmann, Darlene Goldstein, Wolfgang Huber
May 31, 2015 through Jun 05, 2015
http://www.cbg.ethz.ch/news/ascona2015

Transforming Analytical Learning in the Era of Big Data
United States – MI – Ann Arbor
Hosted by: The University of Michigan
Jun 01, 2015 through Jun 26, 2015
http://bigdatasummerinst.sph.umich.edu/

NIMBioS Summer Research Program for Undergraduates and Teachers
United States – TN – Tennessee – Knowxville
Hosted by: National Institute for Mathematical and Biological Synthesis (NIMBioS)
Jun 01, 2015 through Jul 30, 2015
http://nimbios.org/sre/sre2015

REU Site: Modeling and Simulation in Systems Biology (MSSB)
United States – CT – Farmington
Hosted by: UConn Health, Center for Quantitative Medicine
Jun 01, 2015 through Aug 07, 2015
http://cqm.uchc.edu/biomath/

The 15th International Congress of Quantum Chemistry
China – Beijing – Beijing
Hosted by: Tsinghua University & Chinese Chemical Society
Jun 08, 2015 through Jun 13, 2015
http://www.icqc2015.org

2015 MBI Undergraduate Summer Research Program
United States – Ohio – Columbus
Hosted by: Mathematical Biosciences Institute
Jun 08, 2015 through Aug 14, 2015
http://mbi.osu.edu/education/summer-undergraduate-program/

EMBL Conference: The Human Micobiome
Germany – Heidelberg
Hosted by: European Molecular Biology Laboratory
Jun 10, 2015 through Jun 12, 2015
http://www.embl.de/training/events/2015/MET15-01/index.html

Mathematical and Computational Evolutionary Biology
France – Provence – Near Hyeres
Hosted by: Porquerolles Island - Université de Montpellier (France)
Jun 21, 2015 through Jun 25, 2015
http://www.lirmm.fr/mceb2015/

EMBO:EMBL – Enabling Technologiesfor Eukaryotic Synthetic Biology
Germany – Heidelberg
Hosted by: EMBL Heidelberg
Jun 21, 2015 through Jun 23, 2015
http://www.embo-embl-symposia.org/symposia/2015/EES15-04/index.html

Mathematical and Computational Evolutionary Biology
France – Porquerolles
Hosted by: Institut de Biologie Computationnelle, Montpellier
Jun 21, 2015 through Jun 25, 2015
http://www.lirmm.fr/mceb2015/

Biodefense World Summit
United States – MD – Bethesda
Hosted by: Knowledge Foundation, a division of CHI
Jun 22, 2015 through Jun 26, 2015
http://www.biodefenseworldsummit.com

Research Collaboration Workshop for Women in Mathematical Biology
United States – TN – Tennessee – Knoxville
Hosted by: National Institute for Mathematical and Biological Synthesis (NIMBioS)
Jun 22, 2015 through Jun 25, 2015
http://www.nimbios.org/education/WS_wwmb.html

EMBL-EBI-Wellcome Trust: Summer School in Bioinformatics
United Kingdom – Cambridge – Hinxton
Hosted by: Wellcome Trust
Jun 22, 2015 through Jun 26, 2015
https://registration.hinxton.wellcome.ac.uk/display_info.asp?id=495

The qPCR and Digital PCR Congress: USA
United States – CA – San Diego
Hosted by: Global Engage
Jun 25, 2015 through Jun 26, 2015
http://www.globalengage.co.uk/digital-and-qpcr.html

International Synthetic and Systems Biology Summer School SSBSS 2015
Italy – Sicily – Taormina
Hosted by: TaoSciences Research Center
Jul 05, 2015 through Jul 09, 2015
http://www.taosciences.it/ssbss2015/index.html

NIMBios Investigative Workshop: Many-cell System Modeling
United States – TN – Tennessee – Knoxville
Hosted by: National Institute for Mathematical and Biological Synthesis (NIMBioS)
Jul 07, 2015 through Jul 09, 2015
http://nimbios.org/workshops/WS_manycell


13th Annual Conference
ROCKY 2015
ROCKY MOUNTAIN BIOINFORMATICS
December 10 - 12, 2015
Viceroy Hotel
Aspen/Snowmass
Colorado
CONFERENCE CHAIR:
LAWRENCE HUNTER, PH.D.
University of Colorado, Denver
GENERAL INFORMATION: rocky@iscb.org
www.iscb.org/rocky2015
ISCB


GLBIO 2015
GREAT LAKES BIOINFORMATICS CONFERENCE
May 18-20, 2015 Purdue University
West Lafayette, Indiana
www.iscb.org/glbio2015