

# ISCB NEWSLETTER

Quarterly update from the Society and membership at large



## ISMB 2020 COMING TO A COMPUTER NEAR YOU!

ISMB 2020 is headed to a virtual landscape near you! You will now be able to attend live sessions, peruse posters, visit the virtual ISCB Booth and so much more from the comfort of your couch. We will miss Montreal, but ISCB will still bring to the virtual forefront all that makes ISMB 2020 the premier event in the bioinformatics calendar!

Virtual ISMB 2020 will bring together the perfect recipe of cutting-edge science, knowledge building, tutorials, community based and online networking opportunities, student-focused symposium, and so much more for the perfect computational biology enriched online experience. Experience the same science without suffering through the jet lag as all sessions will be made available live as well as on demand to benefit attendees from all time zones.

The virtual conference has the same goal as previous ISMB's and that is to provide an exciting line up of talks, workshops and networking opportunities. The introduction of new COSIs and the continued integration of COSI sessions with the conference and the inclusion of special sessions on new hot topics means participants will be able to indulge in cutting edge science and new technology developments in their field of interest, as well as newly trending topics.

## WHAT'S INSIDE THIS ISSUE:

---

*ISMB 2020 Going Virtual*

*ISCB Congratulates and  
Introduces the 2020 Class of  
Fellows*

*ISMB 2020 Distinguished Keynotes  
& Award Winners*

*ISCBSC Virtual Seminar Series  
Reserve Your Spot at the Virtual  
Career Fair*

*Highlight on ISMB 2020 Special  
Sessions*

*Mark Your Calendars!*

*In Memorium - Those ISCB has  
Lost*

## CONTINUED...

The conference format features proceedings talks and abstract presentations that include highlights (previously published research) and late-breaking research. Each day of the five-day conference includes outstanding keynote lectures, technical tracks, COSI track presentations, a variety of workshops and tutorials, special sessions, equal opportunities activities, a student council day and other focused presentations on other important research area topics in computational biology.

The 2020 COSI track areas are: • 3DSIG • Bio-Ontologies • BIOINFO-CORE • BioVis • BOSC: Bioinformatics Open Source Conference • CAMDA • CompMS • Education • Evolution and Comparative Genomics • Function • HitSeq • MLCSB • MICROBIOME • NetBio • RegSys • iRNA • SysMod • Text Mining • TransMed • VarI • General Computational Biology

The bulk of the ISMB program will be determined by abstract and proceedings sub-missions. Abstract submissions are still being accepted. Submit your work today to be a part of this ground breaking event, bringing you the science you expect directly to your computer or virtual landscape. The Abstract submission deadline is now April 30, 2020. Submit your work today to become a part the first ever virtual ISMB 2020!

**[HTTPS://WWW.ISCB.ORG/ISMB2020-SUBMIT](https://www.iscb.org/ismb2020-submit)**

## ISCB CONGRATULATES THE 2020 CLASS OF FELLOWS

The ISCB Fellows program was created to honor members who have distinguished themselves through outstanding contributions to the fields of computational biology and bioinformatics. During the 2009 inaugural year of the program, the ISCB Board of Directors unanimously conferred Fellows status on the seven winners-to-date of the ISCB Accomplishment by a Senior Scientist Award. 2020 marks the 11th anniversary of the program. Each year since 2010, ISCB has sought Fellows nominations from our members, with eligibility restrictions based on selection criteria focused most heavily on the significance of scientific contributions, and service to our field and to ISCB.

### SERAFIM BATZOGLOU

Stanford University, DNAnexus, Illumina, Insitro, for foundational and innovative contributions to computational genomics in both academia and industry, through his visionary research, mentoring of young leaders and his own leadership and service to the community.

### JUDITH BLAKE

The Jackson Laboratory, for her many contributions to bioinformatics over her distinguished career, most notably the Gene Ontology and the Mouse Genome Databases. These two resources are fundamental pieces of scientific infrastructure serving as foundational tools on which so much of genome-scale science is built.



## MARK BORODOVSKY

Georgia Tech University, for his influential research in developing algorithms of genome analysis as well as recognized leadership in education and community development.



## RITA CASADIO

University of Bologna, as one of the pioneers of machine learning based approaches for predicting protein structures, protein function and the impact of genetic variation. The methods from her group have been highly ranked in international competitions assessing prediction of protein structure (CASP), function (CAFA) and variant impact (CAGI).



## PAUL FLICEK

EMBL-EBI, as a leading figure in the field of genomics and for his tireless efforts advocating for open data and equal opportunities for women in science, including the Earth BioGenome Project in which he advocates workable data sharing practices.



## OSAMU GOTOH

Osamu Gotoh, Kyoto University, a pioneer in computational biology and an exemplary professional researcher, specifically for his contribution to sequence alignment through the development of what is now known as the Smith-Waterman-Gotoh algorithm.



## RAFAEL IRIZARRY

Dana-Farber Cancer Institute / Harvard T.S. Chan School of Public Health, for his pioneering work in expression analysis and development of statistical methods for expression analysis which are some of the most impactful in the entire field. He is one of the leaders of the open-source Bioconductor project since its beginning.



## LAXMI PARIDA

Laxmi Parida, IBM, for her breakthrough contributions to algorithms across diverse applications with impact to human health and society through IBM and clients. She was instrumental in the design of the core engine and accuracy of Watson for Genomics offering in precision oncology.



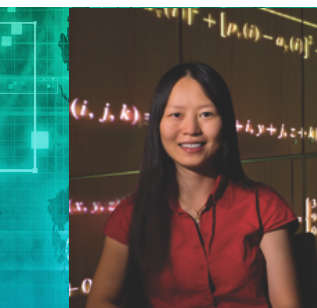
## KATHERINE POLLARD

Gladstone Institutes, University of California, a world leader in developing statistical models and open-source bioinformatics software for biological big data, with an emphasis on genomics. She has uncovered biomedical knowledge that would be missed without her rigorous computational approaches.



## BEN RAPHAEL

Princeton University, as a widely acknowledged leader in the area of computational and mathematical aspects of cancer genomics, with special emphasis on inferring tumor phylogenies.



## ZHIPING WENG

University of Massachusetts Medical School, for her seminal contributions to computational biology and bioinformatics, specifically in structural biology (widely-used ZDOCK suite of protein-protein docking algorithms), regulatory genomics (pipelines for analyzing genomic and epigenomic data from ENCODE and PsychENCODE Consortia) and small-RNA biology (piRNAs).



## XUEGONG ZHANG

Tsinghua University, for the development of computational methods for analyzing transcriptomic and RNA-seq data including the methods DEGseq, DEsingle, SCRL, scFly, and establishing an influential graduate program, educating a generation of students, and promoting the field of computational biology and bioinformatics in China.

---

ISCB will be honoring the 2020 Class of Fellows at the ISCB Town Hall during the Virtual ISMB 2020, July 13-16.

***Congratulations, 2020 Class of ISCB Fellows!***



# HOW YOU CAN SUPPORT ISCB

## Dontate Now!

### ANNA TRAMANTANO FUND

The goal of the Anna Tramontano Fellowship Fund is help reduce the financial burden to the students who are offered these internships by providing travel support. We hope that by providing financial support, reducing costs to the PIs, we will be able to increase the number of internships offered in a given year.



### STUDENT TRAVEL FELLOWSHIP CAMPAIGN

YOU can make a difference in the future of computational biology and bioinformatics by supporting tomorrow's researchers. By donating to ISCB student travel fellowships, you are investing in the future of our science..

### GENERAL RESOURCES FUND

ISCB delivers valuable information about training, education, employment, and relevant news, and provide an influential voice on government and scientific policies that are important to our members and benefit the public. Your membership and generous support helps to make these activities possible.



### SPONSOR MEMBERSHIP DUES OF MEMBERS FROM DEVELOPING COUNTRIES

Your contribution will help those in developing countries who cannot afford membership to join and benefit from ISCB.

## MARK YOUR CALENDARS FOR THESE UPCOMING CONFERENCES

### YOUTH BIOINFORMATICS SYMPOSIUM

May 11, 2020

### VIRTUAL ISMB 2020

July 13 - 16, 2020

### RECOMB 2020

June 22 - 25, 2020

### ECCB 2020

September 05 - 09, 2020

### ISCB-LATIN AMERICA SOIBIO BIONETMX 2020

October 25 - 30. 2020

### RSGDREAM 2020

November 16 - 18, 2020

### ROCKY 2020

December 3 - 5, 2020

### GIW/ISCB-ASIA 2020

December 9 - 11. 2020

<https://www.iscb.org/support-iscb>

# ISMB 2020 DISTINGUISHED KEYNOTES

## ELAINE R. MARDIS

The Steve and Cindy Rasmussen Nationwide Foundation Endowed Chair  
in Genomic Medicine  
Co-Executive Director, The Steve and Cindy Rasmussen Institute for Genomic  
Medicine at Nationwide Children's Hospital  
Professor of Pediatrics  
The Ohio State University College of Medicine  
Columbus, United States

**Presentation Title:** Computational Analysis in Pediatric Cancer Precision Medicine

Elaine Mardis, PhD is co-Executive Director of the Institute for Genomic Medicine at Nationwide Children's Hospital and the Nationwide Foundation Endowed Chair of Genomic Medicine. She also is Professor of Pediatrics at The Ohio State University College of Medicine. Dr. Mardis joined Nationwide Children's Hospital in 2016. Educated at the University of Oklahoma with a B.S. in Zoology and a Ph.D. in Chemistry and Biochemistry, Dr. Mardis did postgraduate work in industry at BioRad Laboratories. She was a member of the faculty of Washington University School of Medicine from 1993-2016. Dr. Mardis has authored over 350 articles in prestigious peer-reviewed journals and has written book chapters for several medical textbooks. She serves as an associate editor for three peer-reviewed journals (Disease Models and Mechanisms, Molecular Cancer Research, and Annals of Oncology) and is Editor-in-Chief of Molecular Case Studies, published by Cold Spring Harbor Press. Dr. Mardis has given lectures at scientific meetings worldwide and was awarded the Morton K Schwartz award from the American Association for Clinical Chemistry in 2016. She has been listed since 2013 as one of the most highly cited researchers in the world by Thompson Reuters. Dr. Mardis has been a member of the American Association for Cancer Research (AACR) since 2007, was the program committee chair for the 2018 AACR Annual Meeting and is the current AACR President. She was elected a Fellow of the AACR Academy, and also was elected to membership in the National Academy of Medicine in 2019.



## LAXMI PARIDA

Master Inventor  
Mathematical Sciences Council  
IBM Academy of Technology  
IBM T. J. Watson Research Center  
New York, United States

Dr. Laxmi Parida is an IBM Fellow, Master Inventor and heads the Computational Genomics at the IBM Thomas J. Watson Research Center, USA. She is a visiting professor at the Courant Institute of Mathematical Sciences, New York. Over the last 10 years, she has led the IBM Science team in the Cacao Consortium (with MARS, USDA), the Genographic Project with National Geographic, the Bioinformatics team in the "Sequence the Food Supply Chain Consortium" across multiple IBM labs in different geographies, and the science team in the personalized cancer medicine system "Watson for Genomics". Her research areas include population genomics, cancer genomics, plant genomics, bioinformatics, algorithms (including AI) and topological data analysis. She has published over 200 peer-reviewed research papers; edited 10 volumes and authored a monograph on pattern discovery in bioinformatics. She holds over 40 US patents. She is on the advisory board of NYU Engineering School and editorial board of BMC Bioinformatics, Journal of Computational Biology and an Associate Editor, IEEE/ACM Transactions on Computational Biology and Bioinformatics and SIAM Journal of Discrete Mathematics. In her spare time she authors a blog and a book on food, and, is a teacher and performer of Argentine Tango.



# ISCB AWARD & ISMB 2020 DISTINGUISHED KEYNOTES

## ISCB INNOVATOR AWARD KEYNOTE XIAOLE SHIRLEY LIU



Biostatistics, Harvard T.H. Chan School of Public Health;  
Co-director, Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute  
United States

The year 2016 marked the launch of the ISCB Innovator Award, which is given to a leading scientist who is within two decades of receiving the PhD degree, has consistently made outstanding contributions to the field, and continues to forge new directions. Xiaole Shirley Liu is the 2020 winner of the ISCB Innovator Award.

Xiaole Shirley Liu has been an innovative and prolific computational cancer biologist. Her research focuses on algorithm development and integrative modeling of high throughput data to understand the specificity and function of regulator genes in tumor development, progression, drug response and resistance. With many contributions to the field, she was a natural choice for this year's award.

In computational biology, her laboratory developed widely used algorithms and tools for transcription factor motif finding, ChIP-chip/seq, chromatin accessibility profiles, CRISPR screen analyses, and tumor immune characterization. Many of her algorithms helped the community adopt new genomics technologies.

In transcription and epigenetic gene regulation, Dr. Liu has been a pioneer in using chromatin dynamics to predict trans-factors and cis-elements involved in biological processes and diseases. As a member of the mod/ENCODE consortium she helped establish best practices in ChIP-chip/seq. She and colleagues generated the first high throughput nucleosome map in the human genome, and identified the chromatin signature of embryonic pluripotency. Her work significantly advanced the understanding of the roles of many genomic and proteomic elements in cancer.

In translational cancer research, Dr. Liu contributed to the discovery of drug response biomarkers, drug resistance mechanisms, and effective combination therapies. Through analyses of large-scale compound and genetic screens as well as tumor profiling cohorts, her group revealed the functions of steroid hormone therapies, epigenetic inhibitors, gamma secretase inhibitor receptor tyrosine kinase inhibitors, and immune checkpoint inhibitors in different cancers.

## ISCB OVERTON PRIZE KEYNOTE JIAN PENG



College of Medicine (by courtesy), Institute of Genomic Biology (affiliate)  
Cancer Center at Illinois (affiliate), National Center of Supercomputing and Applications  
(affiliate)  
University of Illinois at Urbana-Champaign  
United States

The Overton Prize recognizes the research, education, and service accomplishments of early to mid-career scientists who are emerging leaders in computational biology and bioinformatics. The Overton Prize was instituted in 2001 to honor the untimely loss of G. Christian Overton, a leading bioinformatics researcher and a founding member of the ISCB Board of Directors. Jian Peng is being recognized as the 2020 winner of the Overton Prize.

## CONTINUED...

The Overton Prize recognizes the research, education, and service accomplishments of early to mid-career scientists who are emerging leaders in computational biology and bioinformatics. The Overton Prize was instituted in 2001 to honor the untimely loss of G. Christian Overton, a leading bioinformatics researcher and a founding member of the ISCB Board of Directors. Jian Peng is being recognized as the 2020 winner of the Overton Prize.

Jian Peng is an Assistant Professor, Department of Computer Science at the University of Illinois at Urbana-Champaign. Since his undergraduate degree in China, Jian Peng has pioneered the application of deep learning techniques to computational biology, including protein structure prediction, biological network analysis, and drug discovery, by developing highly innovative methods to successfully address non-trivial challenges.

Jian Peng has spearheaded some of the most impressive contributions to our field such as the application of deep learning techniques to protein contact map prediction, his invention of a novel computational framework for simultaneous dimensionality reduction of multiple heterogeneous biological networks, enabling state-of-the-art function prediction and drug discovery and the development of TransposeNet, an approach that translates discoveries from model organisms to human, for which adequate approaches did not previously exist.

Dr. Peng has established himself as internationally-renowned researcher in structure-based, genome scale prediction.

## ISCB ACCOMPLISHMENTS BY A SENIOR SCIENTIST AWARD KEYNOTE STEVEN L. SALZBERG



Bloomberg Distinguished Professor of Biomedical Engineering, Computer Science,  
and Biostatistics Director;  
Center for Computational Biology McKusick-Nathans Institute of Genetic Medicine;  
Johns Hopkins University  
Baltimore, United States

The Senior Scientist Accomplishment Award recognizes a member of the computational biology community who is more than two decades post-degree and has made major contributions to the field of computational biology. Steven L. Salzberg is being honored as the 2020 winner of the ISCB Accomplishments by a Senior Scientist Award.

Steven L. Salzberg, is a Bloomberg Distinguished Professor of Biomedical Engineering, Computer Science, and Biostatistics at Johns Hopkins University, where he also is the Director of the Center for Computational Biology. Dr. Salzberg received his Ph.D in Computer Science, Harvard University in 1989.

His group's research focuses on the development of new computational methods for analysis of DNA from the latest sequencing technologies. Over the years, he and his group have developed and applied software to many problems in gene finding, genome assembly, comparative genomics, evolutionary genomics, and sequencing technology itself, taking new challenges and developing multiple novel, high-impact computational methods, many of which are in extremely broad use.

Dr. Salzberg is one of the most accomplished computational biologists in the world today. Over the past 20 years he has created, over and over again, strikingly innovative computational methods that addressed critical needs in genome analysis. One of his trademarks has been the development of algorithms that are not only accurate but also incredibly efficient, often representing dramatic leaps forward in technology. He has also been a leader in building computational methods to make novel biological discoveries, in species ranging from viruses to bacteria to animals. All of Salzberg's bioinformatics systems have been released as free, open-source software, and he won the 2013 Benjamin Franklin Award for Open Science for his advocacy of open-source software and of open sharing of genome sequence data.



# RECIPIENT OF THE ISCB OUTSTANDING CONTRIBUTIONS AWARD

## JUDITH BLAKE



Professor, The Jackson Laboratory

The Outstanding Contributions to ISCB Award recognizes an ISCB member for outstanding service contributions toward the betterment of ISCB through exemplary leadership, education, and service. The 2020 recipient of the Outstanding Contributions to ISCB Award is Judith Blake.

Judith Blake is a Professor at The Jackson Laboratory. Blake is one of the leaders and founding principal investigators of the Gene Ontology (GO) Consortium, which provides controlled structured vocabularies for molecular biology and has had tremendous impact on both molecular and computational biology. These are fundamental building blocks in a wide variety of bioinformatics and computational applications. Her research is in the area of functional and comparative genome informatics. She develops systems to integrate genetic, genomic and phenotypic information, and has had significant influence on semantic systems use in biology and on bio-curation as a field.

Dr. Blake is a prominent and loyal member of the ISCB, who has served on its Board of Directors and as the ISCB representative to the Federation of American Societies for Experimental Biology. She chairs the ISCB Public Affairs and Policy committee, while performing a variety of other roles on program and review committees.

Dr. Blake is being recognized for her many years of significant contributions to both ISCB and the broad bioinformatics and bio-curation communities.



### RECOMB/ISCB CONFERENCE on REGULATORY & SYSTEMS GENOMICS with DREAM CHALLENGES

NOVEMBER 16-18, 2020 PITTSBURGH, PA



#### DETAILS

**WHEN**  
NOVEMBER 16-18, 2020

**WHERE**  
MELLON INSTITUTE  
CARNEGIE MELLON UNIVERSITY  
4400 FIFTH AVE  
PITTSBURGH, PA 15213

**COST\***

Professional	\$375
Post Doc	\$275
Student	\$250

**Non-Members**

Professional	\$575
Post Doc	\$475
Student	\$450

\*Reflects Early registration rates.  
Rates increase on Oct 13

**MORE INFO**



[www.iscb.org/recomb-regsysgen2020](http://www.iscb.org/recomb-regsysgen2020)

#### SUBMIT TODAY!

**SUBMISSION DEADLINES**

Oral presentations & Posters	September 24, 2020
DREAM Challenge Deadlines	Challenge Specific

**TOPICS:**

- Network visualization and analysis
- Regulatory motifs and modules
- Epigenomics and chromatin state
- Non-coding RNAs
- Regulatory networks
- Co-transcriptional, post-transcriptional, and translational regulation
- Signal transduction networks
- Genetic, molecular, and phenotypic variation and human disease
- Cellular signatures of biological responses and disease states
- Mathematical modeling and simulation of biological systems
- Methods for systematic validation of high-throughput biological predictions
- Single-cell transcriptomics
- Single-cell proteomics
- Metabolomics
- Microbiome
- Machine learning methods for systems biology
- Translational systems biology

**Featuring a Special Session on  
Regulatory and Systems Genomics in Immunology**

We welcome submissions on computational and experimental advances regulatory and systems genomics in immunology. Topics include but are not limited to: immunological regulatory programs and signaling pathways, genetic and non-genetic sources of self/non-self identification, precision epitope identification and targeting, and immune cell regulation.

This special session is sponsored by the Research Center for Cancer Systems Immunology at Memorial Sloan Kettering Cancer Center, an NCI-funded Cancer Systems Biology Consortium (CSBC) Center.

ISCB will present the Accomplishments by a Senior Scientist Award, Overton Prize, Innovator Award and Outstanding Contributions to ISCB Award, at the Virtual ISMB 2020 ([www.iscb.org/ismb2020](http://www.iscb.org/ismb2020)), July 12-16, 2020 where, in addition, Peng, Liu, and Salzberg will present keynote addresses during the conference.

Full bibliographical articles profiling the award recipients will be available in the ISMB 2020 focus issue of the ISCB newsletter later this year, as well as the ISCB Society Pages in OUP Bioinformatics, and F1000 Research ISCB Community Journal.



# YOUTH BIOINFORMATICS CONFERENCE 2020

The International Society for Computational Biology (ISCB) will host the Youth Bioinformatics Symposium (YBS) 2020, exploring computational biology as a **virtual live** event in on **Monday, May 11, 2020**.

YBS is a one-day event that provides an opportunity for students to come together and introduces students to the amazing world of computational biology. Our virtual symposium will allow them to engage with and learn about three popular tools used in research in our online workshops, inform them of the many career areas that bioinformatics is now appearing in, and ask questions.

We hope you join us to make YBS 2020 a success!



**YOUTH BIOINFORMATICS  
SYMPOSIUM**  
Exploring  
Computational  
Biology  
**Monday, May 11, 2020**  
**Live Webinar Event**

**Keynote Speakers**  
**Dr. Judith Blake**  
*The Jackson Laboratory*  
**Dr. Nicola Mulder**  
*University of Cape Town*

**Workshops on:**  
Applications of Deep Learning to Computational Biology Problems  
Analyzing Genetic Data Using the R Programming Language  
High-throughput Sequencing Modeled Through Integrative Genome Viewer  
Introduction to Computational Neuroscience

**More Info:**  
<https://www.iscb.org/ybs2020>

**ISCB**  
INTERNATIONAL SOCIETY FOR  
COMPUTATIONAL BIOLOGY

## VIRTUAL ISMB 2020 SPECIAL SESSIONS

### SST01: SYSTEMS IMMUNOLOGY: COMPUTATIONAL APPROACHES FOR UNDERSTANDING HUMAN IMMUNE SYSTEM AND IMMUNE-RELATED DISEASES

**Organizer(s):** Ferhat Ay, La Jolla Institute for Allergy and Immunology, United States

Duygu Ucar, The Jackson Laboratory for Genomic Medicine, University of Connecticut Health Center, United States

Dysfunctions in the immune system are important contributing factors, if not the driving factors, for a wide range of human diseases. Advances in next generation sequencing techniques are providing us opportunities to closely monitor human immune cells in health and disease, and uncover genomic and epigenomic signatures that can be linked to pathologies. However, the analyses and integration of multi-faceted genomic data are not trivial and require sophisticated computational methods. These methods are developed for data processing and analysis (e.g., imaging, cell sorting, epitope identification, single-cell sequencing), for linking genetic and epigenetic variation to disease susceptibility (GWAS, eQTL, chromatin QTL) and for integrating epigenome, transcriptome and cell abundance measurements from multiple distinct immune cells in the context of immune related diseases such as autoimmune diseases and cancer. As the amount of data collected from individuals to profile their immune systems is increasing, we are encountering a bottleneck to integrate cross-platform data and mine these to test biological hypotheses. This meeting aims to bring together computational scientists who develop methods to effectively analyze single-cell level and bulk measurements from human immune cells in diverse contexts including aging, autoimmune diseases and cancer as well as in the broader context of understanding genetic and epigenetic regulation of gene expression.

### SST02: BIOINFORMATICS OF CORALS

**Organizers:** Lenore Cowen, Department of Computer Science, Tufts University, United States  
Judith Klein-Seetharaman, Department of Chemistry, Colorado School of Mines, United States  
Hollie Putnam, Department of Biological Sciences, University of Rhode Island, United States

Corals are important natural resources that are key to the oceans vast biodiversity and provide economic, cultural, and scientific benefits. As a result of anthropogenic activities, locally and globally, coral reefs are declining rapidly. The environmental sensitivity and symbiotic biological complexity of corals makes understanding the genomic variability that influences vulnerability and resilience of local coral reef systems very challenging. Corals are made up of thousands of different organisms, including the animal host and single celled dinoflagellate algae, bacteria, viruses, and fungi that coexist as a holobiont. Thus, corals are more like cities than individual animals, as they provide factories, housing, restaurants, nurseries, and more for an entire ecosystem, both at the micro and macro levels.

A large amount of genomic, transcriptomic and other omics data from different species of reef building corals, the uni-cellular dinoflagellates, plus coral microbiome data (where corals have possibly the most complex microbiome yet discovered, consisting of over 20,000 different species), is becoming increasingly available for corals. This is a terrific opportunity for bioinformatics researchers and computational biologists to contribute to a timely, compelling and urgent investigation of critical factors that influence reef health and resilience.

We have invited some of the premier experts who are working on bioinformatics of coral reefs to participate in our invited sessions. We will introduce this exciting topic to the ISMB community, with the goal of energizing collaborations and approaches to address the compelling problems in this captivating and complex system. This convergence of data and critical need to address this declining ecosystem provides a timely and impactful topic for ISMB 2020.

### SST03: HUBMAP: INTEGRATING GENOMICS, IMAGING AND MASS-SPECTROMETRY TO CONSTRUCT SINGLE-CELL HUMAN TISSUE MAPS

**Organizers:** Ziv Bar-Joseph, Carnegie Mellon University, United States  
Nils Gehlenborg, Harvard Medical School, United States  
Ajay Pillai, National Human Genome Research Institute, United States

Transformative technologies are enabling the construction of three-dimensional maps of tissues with unprecedented spatial and molecular resolution. Over the next six years, the NIH Common Fund Human Biomolecular Atlas Program (HuBMAP) intends to develop a widely accessible framework for comprehensively mapping the human body at single-cell resolution by supporting technology development, data acquisition, computational analysis and detailed spatial mapping. The HuBMAP project exists within a larger ecosystem of other single cell atlas projects and this session will cover both.

In this special session we will focus on the computational goals, needs and current state of the HuBMAP project. After introducing the types of single cell resolution data HuBMAP proposes to generate a series of speakers will present specific solutions to the following challenges: (1) infrastructure for accessing the data; (2) tools and pipelines for analysis and integration of data; (3) visualizing and querying single cell data at scale; (4) creating a reference common coordinate framework (CCF) to spatially represent single-cell data mapped to specific organs and systems will be addressed.

The intended audience for this special session includes computational researchers that focus on the analysis and integration of large scale biological data. We would both discuss methods, and unmet needs for the analysis of single cell sequence data (scRNA-Seq, scATAC-Seq, etc.), spatial transcriptomics (microscopy) and proteomics (imaging mass spectrometry) data. While the initial pipelines and tools developed for HuBMAP are being implemented by consortium members, we expect that both, additional analysis and additional tools would come from other members of the community who would have access to the data via the web interface and a set of dedicated APIs.



## SST04: SCANGEN: SINGLE-CELL CANCER GENOMICS

**Organizers:** Kieran R Campbell, Lunenfeld-Tanenbaum Research Institute, University of Toronto, & Vector Institute for AI, Canada

The past decade has resulted in technological advances that have given us the unprecedented ability to measure RNA, DNA, and epigenetic modifications at the single-cell level, combined with widespread acquisition of imaging data. This has enabled routine measurement of genomic, morphologic and histopathologic alterations across tens of thousands of cells, discovering new cell types, developmental lineages, and cell-specific mutational patterns. This new data has prompted an explosion in statistical and computational methods development (<http://www.scrna-tools.org/>) with over 540 tools created in the past five years alone.

However, the majority of methods developed remain focused on either technical aspects (such as normalization and differential expression) or on applications in developmental biology such as lineage inference and cell clustering, with relatively little attention applied to the huge potential of single-cell data to unveil the complex biology behind cancer initiation and progression. As one of the first workshops of its kind, this special session will bring together researchers developing computational and statistical methods for single-cell cancer biology. It will focus around (though not be limited to) four core topics:

- **Modeling cancer evolution** As tumors evolve they accumulate both point mutations and large structural rearrangements. The “life-histories” of these tumors are informative of the mutational processes that allow the cancer cells to evade the body’s checkpoints and can be predictive of future evolution and response to therapy. Methods covered under this topic could address: phylogenetic inference from single-cell data; inference of evolutionary processes from single-cell data; identifying singlecell cancer signatures; inference of fitness from single-cell analysis of population dynamics.
- **Integrative analyses of multi-modal and imaging data.** A vast array of measurements can be made at single-cell resolution, including RNA and DNA-sequencing and epigenetic status such as methylation and chromatin accessibility. Methods covered in this topic will include: modelling of joint measurement assays (such as G&T-seq); relating and interpreting measurements from different technologies.
- **Scalable inference at the single-cell level** A typical single-cell RNA or DNA-seq dataset now contains around 100x more cells than it did less than a decade ago. As a result, there is a pressing need for computational and statistical methods that scale to “big data” sizes, particularly since fast computation allows iterative analyses by investigators, aiding biological interpretation. Methods covered in this topic will include: scalable statistical inference for single-cell data using methods such as stochastic optimization; computational tools for dealing with large single-cell datasets.
- **Interactions and perturbations at the single-cell level.** This broad topic concerns methods to understand how cancer cells react to both their environment and external perturbation. Methods could address: how cells interact with their microenvironment; how cells respond to and resist chemotherapeutic interventions; how transcriptional programming and clonal selection are affected by genomic perturbations such as CRISPR.

## SST05: BIOINFORMATICS OUTSIDE THE LAB: HOW TO MOBILIZE ONLINE CITIZEN SCIENTISTS TO ACCELERATE RESEARCH

**Organizers:** Jérôme Waldispühl, McGill University, Canada

This meeting aims to offer a comprehensive overview of the progress and achievements of citizen science applications to biology over the past 10 years. Each talk will discuss a different challenge and describe a solution made possible by the application of crowdsourcing to a specific biological research question.

1. *The origin of scientific games in molecular biology.* This talk will discuss the genesis of the Foldit project and early results that pioneered the field of scientific discovery games for molecular biology. It will provide an overview of the foundation of this field of research.
2. *Building a knowledge base for biologists.* The curation and organization of biological knowledge is an essential resource for researchers, but this task cannot always be fully automated. We will discuss the opportunities offered by crowdsourcing approaches to building and maintaining databases.
3. *Collaborative design of new molecules.* We will discuss how citizen scientists collaborate to discover new scientific principles, and describe algorithmic frameworks allowing to harness the work of crowds.
4. *When gamers meet experimentalists.* Scientific games allow users to process raw experimental data and submit hypothesis to scientists that can be validated in labs. This theme will explore opportunities arising with the creation of new bridges between citizens and scientists.



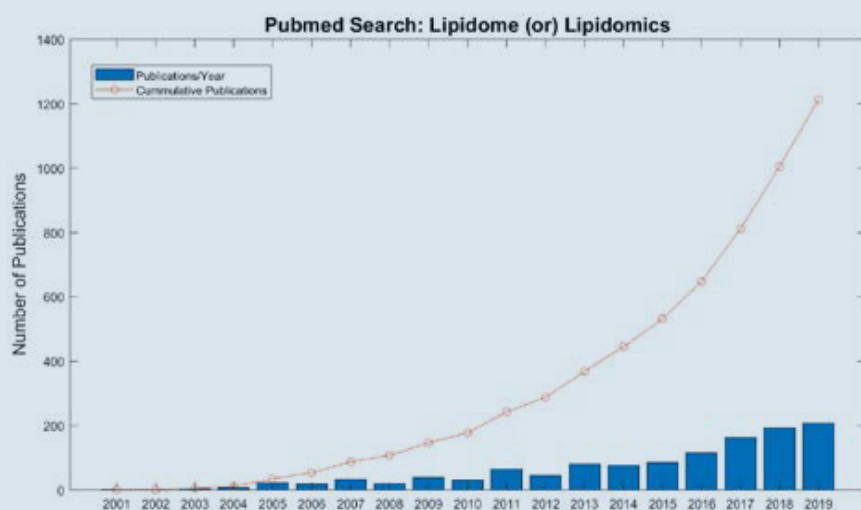
1. *Building communities of citizen scientists.* The success of citizen science initiatives relies on the capacity to build and train large communities of participants. This talk aims to explore strategies to promote engagement and their impact on the performances.
2. *How to turn scientific tasks into casual games.* In this talk, we will show how complex computational tasks can be embedded in casual games in order to make them accessible to broader audiences.
3. *Extreme citizen science: Scaling-up citizen science with video games and AI.* The integration of citizen science tasks in real video games can help to reach very quickly large crowds of participants not naturally exposed to science. Then, the application of AI techniques can also help to exploit the large volume of data collected.

## SST06: INTERPRETING THE LIPIDOME – APPROACHES TO EMBRACE THE COMPLEXITY

**Organizers:** Bobbie-Jo Webb-Robertson, Pacific Northwest National Laboratory, United States  
 Jason McDermott, Pacific Northwest National Laboratory, United States  
 Jeremy Clair, Pacific Northwest National Laboratory, United States

Due to the central role that lipids play in energy metabolism, structure and signaling, current research programs ranging from biomedical to ecological applications have shown an increase in the analysis of lipids [1]. Lipidomics specifically is the large-scale study of the structure and function of all lipids in a sample, termed the lipidome, as well as the interaction of lipids with other biomolecules, including proteins, metabolite and other lipids. Research employing lipidomics is rapidly increasing. As seen in Figure 1, the number of publications referring to lipidomics or the lipidome in their title in the National Library of Medicine PubMed database has increased by over 2-fold in just the last 5 years. This is because recent advancements in mass spectrometry (MS) technology together with bioinformatic developments allow more researchers to perform studies with lipids. However, because the field of lipidomics is younger than the other omics fields and the structure, diversity and complexity of the lipidome differs from than these predecessor omics there is considerable opportunity to advance the field through computation.

The International Lipid Classification, and National Nomenclature Committee referred to as LIPID MAPS (<http://www.lipidmaps.org>) [2] began by categorizing lipids into eight categories based on their chemical and biochemical properties. These classes are further divided by more specific structural and chemical properties. As of January 2018 there are over 40,000 lipid structures in this database and it is expected to continue to grow [3]. This complexity in the lipidome is often under-described by averaging lipids up to the class-based descriptors, losing valuable information about biological function and interactions with other lipids, proteins or metabolites.



Although lipidomics is considered a sub-field of metabolomics it is much earlier in methods development than for metabolite identification and evaluation [4] and the computational needs are still a mystery to many working in bioinformatics and computational biology. While methods and tools have improved over the past few years, computational developments are at the center of the gap in deriving insights from this data. The complexity of lipidomics data means that global-based analyses only identify hundreds of lipids versus the thousands present. In addition, inferring the interaction of the lipidome with the genome and proteome is key to mechanistic understanding.

ISMB participants are interested in cutting edge research and areas of great need for new developments, thus this session will be of interest to a large collection of attendees. To see lipidomics reach its full potential improved strategies for identification and quantification, as well as pathway and integration approaches are needed.

## SST07: THE BRAIN INITIATIVE CELL CENSUS NETWORK

**Organizers:** Jesse Gillis, Cold Spring Harbor Laboratory, United States

The mammalian brain is a complex organ relying on a circuitry involving millions to billions of cells. Identifying neuronal cell types and how these cell types interact is an essential step toward understanding the organization of the brain. The Brain Initiative Cell Census Network (BICCN) aims to define a taxonomy of robust neuronal cell types that can be characterized through multiple modalities, e.g. transcriptional similarity, spatial organization, morphology and connectivity. In the past three years, the BICCN has generated several large reference molecular datasets intended to serve as a broad reference for understanding and analyzing the mammalian brain. As a complex system with an unusual number of modalities to define cell identity, the brain serves as a potential Rosetta stone to clarify many pressing questions regarding the properties and functions of cell types. In this session, we will describe essential properties of the available BICCN data and walk through best practices in its use. As a reference data set on par with the largest previous NIH consortium efforts, the BICCN promises to be an essential resource for future research by data focused scientists in biology.

The BICCN is a very large NIH initiative. While significant outreach is planned (and has been undertaken) to the neuroscience community, with multiple successful symposia by many of the same speakers planned for this special session, there has been little planned outreach to the bioinformatics community. Indeed, with the exception of the Gillis lab, none of the planned speakers have previously attended ISMB. While the analyses from the BICCN are likely to be of broad interest (and currently being submitted as a series of consortium papers, largely available on biorxiv by the time of ISMB2020), its true main contribution is the data generated, available for use by others. For this resource to be fully exploited, the evaluation and characterization by the BICCN needs to be shared with a broader community of data researchers. This special session is intended to make this a reality.

**CONNECTING,  
TRAINING,  
EMPOWERING,  
WORLDWIDE**



[www.iscb.org](http://www.iscb.org)

**Join more than  
3400 researchers  
from around the world.**

- » Leadership, Recognition, Awards, and Travel Funding
- » ISCB Career Center
- » Connection to Communities and Science at your Fingertips
- » Our Collective Voice Promotes, Represents, and Advocates for Science

*Attend just one ISCB official conference and your membership pays for itself.*

**International Society for Computational Biology  
BECOME A MEMBER TODAY!**



# ISCB Academy Webinar Series

In conjunction with the communities of special interest (COSIs), select presentations are invited to give a live-streamed talk about their research.

Access to the webinar series is **complimentary** for ISCB members.

Nonmember can join for a fee.

**SEPTEMBER 30 •  
WEDNESDAY •**

**RAXML-NG: A FAST, SCALABLE AND USER-FRIENDLY  
TOOL FOR MAXIMUM LIKELIHOOD PHYLOGENETIC  
INFERENCE BY ALEXEY KOZLOV, HOSTED BY  
EVOLCOMPGEN COSI**

## OTHER WEBINAR SERIES

Do you want to share your bioinformatics research with a broader audience?

ISCB Student Council webinar series is here to help! Don't forget to submit your abstract before April 10!

Young researchers and students are also encouraged.

iSCB Student *COUNCIL*

# WEBINARS

Computational Biology & Bioinformatics

- Diverse Topics
- All Career Stages
- 20 min + 5 min Q&A
- Shared on YouTube

Submit your abstract here:  
[bit.ly/isbcs\\_webinar\\_submission](http://bit.ly/isbcs_webinar_submission)  
**Deadline: Apr 30, 2020 (23:59 CET)**  
\*presentations will take place between June and August

Follow us on YouTube: [bit.ly/isbcswebinars](http://bit.ly/isbcswebinars)

<https://forms.gle/8PGu7bs8vV8DxJXB9>



JOIN THE COUNTDOWN!

# #3MonthstoGoTilVirtualISMB

Register to attend!

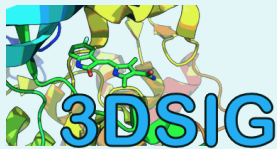


[www.iscb.org/ismb2020-registration](http://www.iscb.org/ismb2020-registration)

## VIRTUAL ISMB 2020 COMMUNITY OF SPECIAL INTEREST (COSI) TRACKS

The organized community sessions (COSI tracks) includes area specific keynote presentations, a selection of talks, which are featured in OUP Bioinformatics in the ISMB 2020 Proceedings supplement, as well as highlight and late-breaking research talks. The 2020 COSI Tracks feature the following communities of special interest:

### 3DSIG: STRUCTURAL BIOINFORMATICS AND COMPUTATIONAL BIOPHYSICS



#### COSI Programme Chairs:

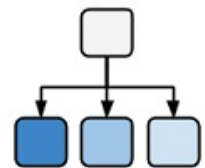
Iris Antes, Technical University of Munich, Germany  
Rafael Najmanovich, Université de Montreal, Canada

It is impossible to fully understand biological systems without understanding the 3D structure of their constituting parts and their interactions. As such the topics relevant for 3DSIG are wide and include, but are not restricted to Structure-based drug discovery including polypharmacology and network pharmacology; Structure representation, classification and prediction;

### BIO-ONTOLOGIES

#### COSI Programme Chairs:

Michel Dumontier, Maastricht University, The Netherlands  
Robert Hoehndorf, King Abdullah University of Science and Technology, Saudi Arabia



Bio-Ontologies Community of Special Interest Group (COSI) covers the latest and most innovative research in the application of ontologies, the organisation and dissemination of knowledge, and the development and application of knowledge-based methods in biomedicine and life sciences.



## BIOINFO-CORE



### COSI Programme Chairs:

Madeline Gogol, Stowers Institute,  
United States

Rodrigo Ortega Polo, Agriculture and  
Agri-Food Canada

Alberto Riva, University of Florida,  
United States

Bioinfo-core is a worldwide body of people that manage or staff bioinformatics cores within organizations of all types including academia, academic medical centers, medical schools, biotech and pharma.

\*Accepting submissions for posters only\*

## CAMDA: CRITICAL ASSESSMENT OF MASSIVE DATA ANALYSIS



### COSI Programme Chairs:

David Kreil, Boku University Vienna,  
Austria

Joaquin Dopazo, Fundación  
Progreso y Salud, Spain

Paweł P Łabaj, Austrian Academy of  
Sciences, and Jagiellonian  
University, Poland

Wenzhong Xiao, Harvard Medical  
School, United States

The large, complex data sets for the Critical Assessment of Massive Data Analysis (CAMDA) contest include built-in truths for calibration. In an open-ended competition, however, both seasoned researchers and cunning students push the boundaries of our field, with unexpected questions or angles of approach often bringing the most impressive advances. Join and celebrate with us CAMDA 2000-2020!

## EDUCATION: COMPUTATIONAL BIOLOGY EDUCATION



### COSI Programme Chairs:

Annette McGrath, CSIRO, Australia  
Patricia M. Palagi, SIB Swiss

Institute of Bioinformatics,  
Switzerland

Education-COSI focuses on bioinformatics and computational biology education and training across the life sciences.

## BIOVIS: BIOLOGICAL DATA VISUALIZATION



### COSI Programme Chairs:

Danielle Albers, University of  
Colorado at Boulder, United

StatesThomas Höllt, Leiden  
University Medical Center, The  
Netherlands

Michael Krone, University of  
Tübingen, Germany

The BioVis track aims to educate, inspire, and engage bioinformatics and biology researchers in state-of-the-art visualization research and visualization researchers in problems in biological data visualization.

## COMPMS: COMPUTATIONAL MASS SPECTROMETRY



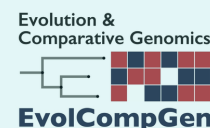
### COSI Programme Chairs:

Olga Vitek, Northeastern University,  
United States

Wout Bittremieux, University of  
California San Diego, United States

COSI CompMS promotes the efficient, high-quality analysis of mass spectrometry data through dissemination and training in existing approaches and coordination of new, innovative approaches.

## EVOLUTION AND COMPARATIVE GENOMICS



### COSI Programme Chairs:

Lars Arvestad, Stockholm University,  
Sweden

Edward L. Braun, University of  
Florida, United States

Evolution and comparative genomics are deeply intertwined with computational biology. Computational evolutionary methods, such as phylogenetic inference methods or multiple sequence alignment are widely used, yet remain far from "solved" and are indeed intense areas of research.

## FUNCTION INCORPORATING CAFA 4: GENE AND PROTEIN FUNCTION ANNOTATION



### COSI Programme Chairs:

Iddo Friedberg, Iowa State  
University, United States  
Kim Reynolds, University of Texas  
Southwestern Medical Center, United  
States  
Mark Wass, University of Kent,  
United Kingdom

The mission of the Function Community of Special Interest (Function-COSI) is to bring together computational biologists, experimental biologists, biocurators, and others who are dealing with the important problem of gene and gene product function prediction, to share ideas and create collaborations.

## MICROBIOME



### COSI Programme Chairs:

Aaron Darling, University of  
Technology Sydney, Australia

The MICROBIOME Community of Special Interest aims at the advancement and evaluation of computational methods in microbiome research, especially metaomic approaches. Based on the Critical Assessment of Metagenome Interpretation (CAMI), the COSI supplies users and developers with exhaustive quantitative data about the performance of methods in relevant scenarios.

## NETBIO: NETWORK BIOLOGY



### COSI Programme Chairs:

Martina (Tina) Kutmon, Maastricht  
University, Netherlands

As large scale, systems-level data are becoming increasingly available, modeling and analyzing them as networks is widespread. Network Biology Community serves to introduce novel methods and tools, identify best practices and highlight the latest research in the growing and interdisciplinary field of network biology.

## HITSEQ: HIGH- THROUGHPUT SEQUENCING



### COSI Programme Chairs:

Can Alkan, Bilkent University, Turkey  
Ana Conesa, University of Florida,  
United States  
Francisco M. De La Vega, Stanford  
University, United States  
Dirk Evers, Dr. Dirk Evers  
Consulting, Germany  
Gang Fang, Mount Sinai School of  
Medicine, United States  
Kjong Lehmann, ETH-Zürich,  
Switzerland  
Layla Oesper, Carleton College,  
United States

HiTSeq is a community of special interest devoted to the latest advances in computational techniques for the analysis of high-throughput sequencing (HTS) data. Sessions will be devoted to discussing the latest advances in computational techniques for the analysis of high-throughput sequencing (HTS) datasets and will provide a forum for in-depth presentations of the methods and discussions among the academic and industry scientists working in this field.

## MLCSB: MACHINE LEARNING IN COMPUTATIONAL AND SYSTEMS BIOLOGY



### COSI Programme Chairs:

Christoph Lippert, University of  
Potsdam, Germany  
Sara Mostafavi, University of British  
Columbia, Canada

Systems Biology and Machine Learning meet in the MLCSB COSI. The community is the place for researchers of these areas to exchange ideas, interact and collaborate.

## REGSYS: REGULATORY AND SYSTEMS GENOMICS



### COSI Programme Chairs:

Ferhat Ay, La Jolla Institute, United States

Ziv Bar-Joseph, Carnegie Mellon University, United States

Anaïs Bardet, CNRS - Université de Strasbourg, France

Mathieu Blanchette, McGill University, Canada

Raluca Gordan, Duke University, United States

Shaun Mahony, Penn State University, United States

Anthony Mathelier, University of Oslo, Norway

Alejandra Medina-Rivera, National University of Mexico

Judith Zaugg, EMBL, Germany

Regulatory genomics involves the study of the genomic control system, which determines how, when and where to activate the blueprint encoded in the genome. Regulatory genomics is the topic of much research activity worldwide. Since computational methods are important in the study of gene regulation, the RegSys COSI meeting focuses on bioinformatics for regulatory genomics.

## SYSMOD: COMPUTATIONAL MODELING OF BIOLOGICAL SYSTEMS



### COSI Programme Chairs:

Laurence Calzone, Institut Curie, France

Andreas Dräger, University of Tübingen, Germany

Juilee Thakar, University of Rochester Medical Center, United States

The Computational Modeling of Biological Systems (SysMod) aims to create a forum for systems modelers and bioinformaticians to discuss common research questions and methods. The session will focus on the conjoint use of mathematical modeling and bioinformatics to understand biological systems functions and dysfunctions.

## IRNA: INTEGRATIVE RNA BIOLOGY



### COSI Programme Chairs:

Yoseph Barash, University of Pennsylvania, United States

Klemens Hertel, UC Irvine, United States

Michele Scott, University of Sherbrooke, Canada

iRNA track covers the full range of research topics in the field of RNA Biology, from computational and high-throughput experimental methods development to their application in different aspects of RNA processing, structure, and function.

## TEXT MINING: TEXT MINING IN BIOINFORMATICS



NEW IN 2020

### COSI Programme Chairs:

Cecilia Arighi, University of Delaware, United States

Lars Juhl Jensen, University of Copenhagen, Denmark

Robert Leaman, NCBI/NLM/NIH, United States

Zhiyong Lu, NCBI/NLM/NIH, United States

This session brings together researchers that create text mining tools with researchers who currently use or are interested in using text mining tools to make new discoveries. The primary goal is to link at least two distinct audiences: those who are not text mining specialists, but who could use the results in their work (e.g., bioinformaticians and computational biologists).

## TRANSMED: TRANSLATIONAL MEDICAL INFORMATICS



### COSI Programme Chairs:

Wei Gu, University of Luxembourg

Venkata Satagopam, University of Luxembourg

Mansoor Saqi, Kings College London, United Kingdom

Maria Secrier, University College London, United Kingdom

TransMed covers the current developments in the field of clinical and translational medicine informatics. Analysis of large amounts of multi-omics, imaging (medical and molecular), mobile sensor, clinical and health records data is paving the way for precision medicine. In the TransMed track, we will explore the current status of computational biology and advance machine learning approaches within the field of clinical and translational medicine.

## VARI: VARIANT INTERPRETATION



### COSI Programme Chairs:

Emidio Capriotti, University of

Bologna, Italy

Hannah Carter, University of

California, San Diego, United States

Antonio Rausell, Imagine Institute for

Genetic Diseases, France

The VarI COSI meeting is dedicated to the recent advances in the analysis and interpretation of the genetic variants

## GENERAL COMPUTATIONAL BIOLOGY

### COSI Programme Chairs:

Virginie Uhlmann, European

Bioinformatics Institute (EMBL-EBI),

United Kingdom

Xuegong Zhang, Tsinghua

University, China

Novel techniques in emerging areas of computational biology, including intersections with other fields.



28TH CONFERENCE ON

**Intelligent Systems  
for Molecular Biology**

JULY 13-16, 2020



## DON'T MISS YOUR CHANCE TO SUBMIT

Abstracts

April 30, 2020

COVID-19 Abstracts

May 14, 2020

**NEW IN 2020**

CAMDA Abstracts

May 25, 2020

Late Breaking Posters

May 28, 2020

<https://www.iscb.org/ismb2020-submit>



# ISCB IN MEMORIAM...

## CYRUS CHOTHIA

---

Cyrus Chothia, who died in November 2019 was a renowned biochemist whose work focused on protein structures, their function and evolution. Chothia was a household name within the ISCB community, having been awarded the ISCB Senior Scientist Award in 2015 in honor of his work in using computational methods to understanding protein structure.



\*Chothia was deeply interested in history as a young student in England but was discouraged from pursuing this subject by teachers. "Because my foreign languages were so bad," Chothia recalled, "my teachers said I couldn't be an historian. I turned to chemistry." As a youth, Chothia remembers being fascinated by a hugely popular BBC television series by Nobel Laureate John Kendrew called "The Thread of Life." He said, "I was enthralled by the series. I knew what I wanted to do, I wanted to be a molecular biologist and go Cambridge. I did it eventually in a roundabout way." Michael Levitt, a 2014 Nobel Laureate and a colleague of Chothia's, was also deeply moved by Kendrew's series as a young man.

Chothia completed his Bachelor of Science degree at Durham University in 1964 and his Master of Science degree at Birkbeck College, University of London, in 1967. He then pursued his PhD under the guidance of Peter Pauling at University College London. Chothia has fond memories of his time under Pauling's supervision. "He was a generous supervisor and we got along extremely well," he said. His PhD research examined the conformations of molecules at nerve receptors. This marked the beginning of his lifelong research interest in protein structures.

After his PhD, Chothia went to the LMB for postdoctoral training, during which time he got to know Joel Janin and Michael Levitt. His experience at LMB was quite different from his time under Pauling, and he and his new supervisor parted after three years. But this presented Chothia with an opportunity to travel abroad to different labs. "[Levitt and Janin] helped me find new places. Kendrew helped me get an EMBO fellowship to come to America to Fred Richards' lab at Yale. America was marvelous scientifically. People were excited about their work. I then went to the Weizmann Institute in Israel for six months and then to Paris to Joel's laboratory for two years. This was my 'Grand Tour' and it was actually very valuable. I went to all these different labs and learned many things that I could use when I got back to Cambridge." This was a very fruitful period for Chothia, and he published numerous papers with Levitt and Janin. Chothia and Levitt developed the "all-a, all-b, a/b and a+b" classification of protein structures, and Chothia and Janin worked out the underlying principles required for protein-protein recognition and packing of protein secondary structures.

In 1976, Chothia came back to England and was affiliated with the University College London and the LMB. He was named the E.P.A. Cephalosporin Fund Senior Research Fellow of the Royal Society in 1980, which offered him stable funding to work as an independent researcher for ten years. Chothia recollected, "During that time, I met Arthur Lesk and we got on enormously well. Lesk was there for more than ten years. Working with him and others was very productive." Their work covered a wide range of topics in structural biology, including how protein structures change to adapt to mutations, mechanisms by which proteins can transmit information to distant sites in a structure, and the observation that a small repertoire of structures exist for the main chain conformation of immunoglobulin hypervariable regions, and these structures can be predicted from the amino acid sequences. By 1990, Chothia recalled how his lab work changed. "I was given a permanent position at Cambridge and begin to have students," he said. "That made things somewhat different and I liked working with them. Before this, I worked with my contemporaries, and they would tell me when I was talking nonsense. Joel Janin and I used to have fierce arguments. When I started taking students the relationship became somewhat different: it was important to explain the terms of the argument."

In 1992, Chothia's work on protein structures led him to propose that most proteins are comprised of domains that come from a limited number of families. Together with Tim Hubbard, Alexei Murzin, and Steven Brenner, he created the Structural Classification of Proteins (SCOP) database. SCOP contains all entries in the Protein Data Bank and provides a detailed and comprehensive description of the structural and evolutionary relationships between all proteins whose structure is known.[i] More recently, Chothia developed the SUPERFAMILY database with Julian Gough, which is a structural and functional annotation for all proteins and genomes.[ii] The SUPERFAMILY annotation is based on a collection of hidden Markov models that represent structural protein domains at the SCOP superfamily level. Superfamilies group together domains with an evolutionary relationship.

---

[1] Hubbard T, Murzin A, Brenner S, Chothia C. 1997. SCOP: A structural classification of proteins database. Nucleic Acids Res. 25(1):236-239.[1] Gough J, Chothia C. 2002. SUPERFAMILY: HMMs representing all proteins of known structure. SCOP sequence searches, alignments and genome assignments. Nucleic Acids Res. 30(1):268-272. \*2015 ISCB Senior Scientist Award Article

---

## RUSS DOOLITTLE

---

Russell Doolittle, considered an illustrious colleague and treasured faculty member, who made extraordinary contributions to the field of molecular evolution and the study of the clotting of blood, died Friday, October 11, 2019, from complications of metastatic melanoma. He was 88 years old.



Doolittle was born January 10, 1931, in New Haven, Conn. He attended Wesleyan University from 1948 to 1952, served in the U.S. Army, arriving in Korea a few days before the armistice in July of 1953, and married Frances Tynan, his wife of 64 years, in 1955. He earned a master's degree in education from Trinity College in Hartford, Conn., in 1957 while teaching science at New Milford High School. At the end of that year, he entered the doctorate program of the Department of Biological Chemistry in the Harvard Medical School, where he performed his doctoral research in the laboratory of J. Lawrence Oncley. After graduating in 1961, he spent a year teaching at Amherst College. Then, in 1962, he joined the laboratory of Birger Blombäck at the Karolinska Institutet in Sweden, where he spent two years on a postdoctoral fellowship from the National Institutes of Health. In 1964, he joined the laboratory of S. Jonathan Singer as an assistant research biologist in the Department of Biology at UC San Diego. In 1965, he accepted the position of assistant professor in what is now the Department of Chemistry and Biochemistry and, in 1987, he was also appointed in what is now the Division of Biological Sciences, ultimately becoming a member of the Section of Molecular Biology.

At UC San Diego, Doolittle studied the amino acid sequences of fibrinopeptides. His work in Sweden showed that the differences between mammal species in the amino acid sequences of these short peptides provided evidence of evolutionary relationships, so he set up a collaboration with the San Diego Zoo to obtain blood drawn from animals. This led to an article in a local periodical about "Dr. Doolittle" at the zoo.

Doolittle also realized in Sweden that the rate at which the sequence of amino acids in fibrinopeptides changed with time provided information about the period of time over which the separation of species within the different orders of animals occurred. He used this knowledge to create phylogenies for the cetartiodactyla order (pigs, camels, llamas, mule deer, reindeer, red deer, cape buffalo, bison, goats, and sheep), as well as for the primate order (macaque, green monkey, baboon, drill, gibbon, chimpanzee, and man). These were seminal studies in the nascent field of molecular evolution, which has exploded over the last 50 years.

Doolittle's study of molecular evolution resulted in one of the first online compendiums of all the amino acid sequences available at the time-a forerunner of the massive data banks now available online. He also began studies of the amino acid sequence of the complete molecule of fibrinogen from the lamprey, the member of the phylum of chordates most distant from mammals. With these results, he extended his quest for the ancestors of fibrinogen to invertebrates. In 1996, he published a phylogeny that included the divergence between plants, fungi, and animals dating from a billion years ago. He also advised others who were doing similar work around the world.

In the 1970s, while studying fibrinogen-the molecule in the blood that produces the clots that staunch bleeding but also can accumulate in the wrong places, causing strokes and heart attacks-Doolittle determined the sequences of the amino acids in the three polypeptides that comprise human fibrinogen, 1,810 amino acids in all. This was a monumental undertaking that tested the limits of existing techniques at the time.

Eager to do more, Doolittle deduced the structure of a molecule of fibrinogen and proposed a detailed molecular mechanism as to how the fibrinogen could polymerize to cause the blood to clot. Validation of his mechanism required an atomic structure of fibrinogen, which up to that time had defied the techniques of X-ray crystallography. Undaunted, Doolittle learned the necessary methods and persevered until he obtained several atomic structures, including that of the entire molecule of fibrinogen, thus proving his molecular explanation for blood clotting.

One of Doolittle's most important and broadly relevant contributions was his work on comparative protein sequence and structure. He also provided key insights into the evolutionary conservation of genes and the nature of gene families. He established the first building blocks of the human genome project, elevating the understanding of biology to a gene level in the 1970s and 1980s, setting the stage for future discoveries. He was a true pioneer in early gene science.

For the elegance and importance of his work, Doolittle was elected a member of the National Academy of Science in 1984. He received a Guggenheim Fellowship in 1984, the Paul Ehrlich Prize in 1989, and the John J. Carty Award for the Advancement of Science in 2006. Over the years, he presented 18 named lectures to various university faculties and scientific societies. He was a hands-on scientist and mentor, who also served UC San Diego at large as chair of the Department of Chemistry and Biochemistry, playing a role in establishing the basic science programs for the undergraduate curriculum in the department, as well as for the developing School of Medicine. He also served on 25 committees of the Academic Senate, as chairman of the Academic Senate, chair of the Executive and Policy Committee, and member of the university-wide Academic Council.

Doolittle, who was a marathon runner during his younger years and even made a run for Congress, was known to have high standards for scientific research, a willingness to do the work to meet those standards, and expectations of the same from his peers. He was an entertaining raconteur with many amusing stories, and he was a willing mentor to younger scientists. He always had a good word to say about his colleagues, and he was an enduring friend.

Doolittle was devoted to his wife, Fran, and their sons, Larry and Will, who survive him. He is also survived by his four grandchildren, who range in age from 16 to 30, and his siblings Donald Doolittle and Kathy Gunther, who both reside in Connecticut.

Donations can be made in Russell Doolittle's memory, by making checks payable to UC San Diego Foundation for the Russell Doolittle Scholarship Endowment, Fund #F-6651-6651. Please mail checks to UC San Diego Gift Processing, 9500 Gilman Drive, #0940, La Jolla, CA 92093-0940. Donations can also be made online at <https://espi.ucsd.edu/make-a-gift>; type 6651 in the Search for More Giving Options field. Arrangements for a celebration of life ceremony are pending. Read Doolittle's full obituary at <https://chemistry.ucsd.edu/news/doolittle.html>.

---

[https://biology.ucsd.edu/about/news/article\\_110619.html](https://biology.ucsd.edu/about/news/article_110619.html)

## JAMES TAYLOR

---

*Good ideas don't have owners—they belong to everyone* (@jxtx) James Taylor started his professional path at the University of Vermont, where he received a BS in Computer Science in 2000. In 2003, after working as a software engineer in the private sector, he found that his real purpose in life was elsewhere. That year he began his graduate studies in Computer Science at Penn State and joined the nascent Center for Comparative Genomics and Bioinformatics. He used to refer to the Center as 'BX'—a contraction of BioinformatiX. The name stuck; 'BX' is still a part of many URLs (e.g. the original Galaxy URL at <http://g2.bx.psu.edu>) and is the origin for the name of the bx-python software package (<https://github.com/bxlab/bx-python>).

The interdisciplinary faculty of the Center included Francesca Chiaromonte, Ross Hardison, Kateryna Makova, Webb Miller, and Anton Nekrutenko. This group became the core of James' academic family, a family that has now lost its most brilliant and prolific child.



# ISCB CAREER CENTER

## ENGAGE INFORM RECRUIT

*REPRESENTING ORGANIZATIONS WORLDWIDE  
START YOUR JOB SEARCH HERE!*

### FIND A JOB

**IISC MEMBERS HAVE COMPLIMENTARY ACCESS TO THE ISCB JOBS BOARD**

### POST A CV:

**ISCB MEMBERS GET FREE UPLOADS**

**NONMEMBERS MAY PARTICIPATE FOR \$100 PER CV**

### POST A JOB:

**ISCB MEMBERS RECEIVE COMPLIMENTARY POSTINGS AND CV REPOSITORY ACCESS**

**NONMEMBERS CAN REGISTER & POST FOR \$300 PER JOB POSTING**

*LOOKING FOR A STRONG JOB CANDIDATE?  
ISCB'S CAREER CENTER IS THE PLACE TO START!*

[WWW.ISCB.ORG/CAREERCENTER](http://WWW.ISCB.ORG/CAREERCENTER)

James concluded his PhD work under the supervision of Webb Miller and Francesca Chiaromonte in just three years—during which he blossomed as a methods developer and as a scientist. He published extensively on comparative genome analysis, gene regulation and molecular evolution—a total of eleven papers 1–11 which were hugely instrumental in advancing the research of the Center. Ross Hardison, former head of the Center, remembers:

I may have been on James' thesis committee, but I learned far more from him than he did from me. In those days, the major data sources for functional genomics were alignments of genome sequences from different species—a specialty of Webb Miller. James was an early developer of machine learning methods to find signals in those multi-species alignments that were predictive of gene regulatory regions. He realized that those methods required substantial dimensional reduction—a specialty of Francesca Chiaromonte—to be effective. He had an amazing command of the statistical, computational, and biological frameworks in which he was working. With his high energy and creativity, he generated effective, publicly released predictions of regulatory elements by 2006. An enduring highlight of my career was my experimental lab testing many of those predictions, and finding that an impressive portion of them did affect levels of gene expression in transfected cells.

James and I continued to work together on many projects, always with the goal of finding out what is important in our genomes. It is hard to think of continuing on without him.

Midway through his graduate studies James became entangled in a messy project that was given the name “GalaXY”. This name was coined by Robert Harris, a fellow graduate student and LASTZ 12 developer, by adding ‘X’ and ‘Y’ to the name of Galaxy's predecessor, GALA 13. GALA was the brainchild of Ross Hardison and Webb Miller, who wanted to link genome alignments and annotations and provide tools for operating on those data. They convinced James and Anton Nekrutenko to take it on as a potentially interesting side project. It is hard to pinpoint the exact time when James and Anton started working on Galaxy together, but most likely it was in the middle of a long night of random pub traversals in Glasgow, Scotland, during the 2004 annual meeting of the International Society for Computational Biology (ISCB).

After graduating from Penn State, James gained post-doctoral experience as a visiting member of the Courant Institute of Mathematical Sciences at the New York University (2006 - 2008). During these two years, and ever since, Galaxy became one of his principal projects. Its now iconic features such as the three-pane interface, the “noodly” workflow editor, and the dynamic genome browser Trackster, were built by James while at NYU.

James' move to Emory University as an assistant professor in 2008 coincided with the rapid emergence of cloud computing, and it became obvious that the future of Galaxy was cloud-dependent. Together with his new team—Enis Afgan, Dannon Baker, Kanwei Li, and Jeremy Goecks—James quickly adapted Galaxy to cloud infrastructure, making it the first comprehensive data analysis resource in Life Sciences to cross this bridge 14,15.



From the very beginning, James' vision for the future of Galaxy was about creating and supporting a community of developers and users. This took shape with the Galaxy Developer Conference held at Cold Spring Harbor Laboratory in the summer of 2010, which has since become a yearly event that draws more than 200 attendees. He believed that to be successful and, most importantly, useful, a project must be open and cannot be owned by a single lab or a PI 16. He was firmly dedicated to this idea. Today Galaxy is not associated with a particular group or an institution. It is a community effort supported by hundreds of developers worldwide. This, perhaps more than anything else, ensures that James' legacy will endure.

James once said that he would change his middle name to Reproducibility if it would help move the cause forward. He was an advocate for reproducible science long before it became fashionable, and he made it a central tenet of Galaxy. He could be downright evangelical about it. Martin Morgan of Bioconductor, a collaborator on the AnVIL Project, shared:

We in the Bioconductor community knew James as a Galaxy project leader, a strong advocate for open and reproducible science, and an enthusiastic and inspirational colleague. James recognized Bioconductor as a kindred spirit, and bridged the relationship between Galaxy and Bioconductor through his leadership of the NHGRI AnVIL project.

James' leadership has had deep consequences for the way Bioconductor now navigates toward large-scale cloud-based computing.

James was particularly concerned with the inadequate state of quantitative education in life sciences. He developed many of the early training materials that became the foundation of the Galaxy Training Network 17. He continued to make important contributions to biomedical education after he joined the Biology Department at Johns Hopkins University (JHU) in 2014. At JHU he participated in the development of massive open online courses (MOOCs), and he taught data analysis and scientific computing to biologists.

Vince Hilser, chair of the Johns Hopkins Biology department, describes James as a bedrock of the department:

He came in 2014, and it was transformational. He was this catalyst for change, with a huge positive impact. His presence in the department opened up many areas for research, as he was able to help other faculty members uncover new insights by revealing similarities between the proteins they were studying and those in other organisms.

People who worked for James thought the world of him, and many people in his group followed him from Emory to Johns Hopkins, and are still in his group today. When a new employee asked him and Anton whether it was "better to ask for permission or forgiveness" James's response was "Don't let us be a bottleneck." That was typical—always thinking of a third option where everyone else only saw two. Enis Afgan, in his memoriam, describes James:

He ran his lab not as a boss but as a friend and a colleague that inspired each individual to do their best work. He was an enabler. He was someone you tried your best not to disappoint.

While we celebrate James' transformative vision of computational biology and bioinformatics, he was also just a lot of fun to be with. He had a finely tuned palate for good beer, he was an energetic and inquisitive conversationalist, and he had an uncanny ability to stay up late talking about science, politics, and all things related, but still be at his best the next morning.

# ISMB 2020 VIRTUAL CAREER FAIR

*Within the ISMB 2020 virtual conference environment, ISCB is pleased to offer the 2020 Career Fair!*

## SIGN UP TODAY:

- 5-MINUTE PRESENTATION DURING THE VIRTUAL CAREER FAIR\*
- ORGANIZATION DESCRIPTION POSTED WITHIN THE VIRTUAL CAREER CENTER
- ONE-ON-ONE APPOINTMENT SCHEDULER FOR THE DURATION OF THE VIRTUAL CONFERENCE
- ISMB 2020 JOBS BOARD POSTINGS (UP TO FIVE POSTINGS)
- JOBS MIGRATED TO ISCB CAREER CENTER FOR 60 DAYS AFTER THE CONFERENCE

*Exclusive networking event held online for recruiters to find that perfect candidate.*

**ISCB MEMBER 145 USD**  
**NONMEMBER 455 USD**

[HTTPS://WWW.ISCB.ORG/ISMB2020-CAREERCENTER](https://www.iscb.org/ismb2020-careercenter)



James met his future wife, Meredith Greif, when they were graduate students at Penn State. They were married in 2009. They have pursued their careers in a mutually supportive manner, him holding a faculty appointment at Emory University, and Meredith at Georgia State University. Both then moved to Johns Hopkins University. We extend our deepest sympathy to her on this shattering loss.

This memoriam can only hint at how remarkable James was as a human being. We could not capture who he was even if we were given volumes to fill. Comments have come in (and continue to come in) from all over the world, and they touch on both his humanity and his intellect. Melissa Wilson, an associate professor at Arizona State and a fellow PhD student, captures this perfectly:

He is leaving a hole that cannot be filled. I met James when I started grad school at Penn State. I had no idea what to expect. We were both students, but he was clearly in a league of his own. James was one of the smartest people I've ever known, in a way that wasn't (as) intimidating. Talking with James was easy and enlightening. He advocated and put in the work for better science, in all aspects. James supported my science and my career in ways I don't think I can ever fully communicate. As a colleague, he believed in me. As a friend, he expected me to do good. He supported my growth, selflessly. Most of all, James was an advocate. He advocated for his students, for trainees, for open science, for open data, for making this world a better place.

James hated self-promotion. For him, the biggest human flaw was baseless arrogance. He believed that everything should be based on merit, which for him meant top-notch science, technical perfection, and a healthy dose of idealism. After all, he was the brightest, most unapologetic, and kindest idealist that the Galaxy project, his students, and his many colleagues and friends will ever know. This post from Karen Reddy, a colleague at Johns Hopkins, sums it all up:

James was, simply put, an amazing person. He was certainly an amazing scientist, but so much more than that. He was what we all hope to be as scientists, mentors, colleagues, and friends. We will miss you James.

Enis Afgan, Dannon Baker, Francesca Chiaromonte, Dave Clements, Nate Coraor, Jeremy Goecks, Ross Hardison, Anton Nekrutenko and all of Galaxy Team and Galaxy Community.

April 5, 2020

- 
1. Gibbs, R. A. et al. *Genome sequence of the Brown Norway rat yields insights into mammalian evolution*. *Nature* 428, 493–521 (2004).
  2. Kolbe, D. et al. *Regulatory potential scores from genome-wide three-way alignments of human, mouse, and rat*. *Genome Res.* 14, 700–707 (2004).
  3. Hillier, L. W. et al. *Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution*. *Nature* 423, (2014).
  4. Taylor, J. *Clues to function in gene deserts*. *Trends Biotechnol.* 23, 269–271 (2005).
  5. King, D. C. et al. *Evaluation of regulatory potential and conservation scores for detecting cis-regulatory modules in aligned mammalian genome sequences*. *Genome Res.* 15, 1051–1060 (2005).
  6. Chimpanzee Sequencing and Analysis Consortium. *Initial sequence of the chimpanzee genome and comparison with the human genome*. *Nature* 437, 69–87 (2005).
  7. Hillier, L. W. et al. *Generation and annotation of the DNA sequences of human chromosomes 2 and 4*. *Nature* 434, 724–731 (2005).
  8. Giardine, B. et al. *Galaxy: A platform for interactive large-scale genome analysis*. *Genome Res.* 15, (2005).
  9. Wang, H. et al. *Experimental validation of predicted mammalian erythroid cis-regulatory modules*. *Genome Res.* 16, 1480–1492 (2006).
  10. Taylor, J., Tyekucheva, S., Zody, M., Chiaromonte, F. & Makova, K. D. *Strong and weak male mutation bias at different sites in the primate genomes: insights from the human-chimpanzee comparison*. *Mol. Biol. Evol.* 23, 565–573 (2006).
  11. Taylor, J. et al. *ESPERR: learning strong and weak signals in genomic sequence alignments to identify functional elements*. *Genome Res.* 16, 1596–1604 (2006).
  12. Harris, B. *lastz*. (Github).
  13. Giardine, B. et al. *GALA, a database for genomic sequence alignments and annotations*. *Genome Res.* 13, 732–741 (2003).
  14. Afgan, E. et al. *Galaxy CloudMan: delivering cloud compute clusters*. *BMC Bioinformatics* 11 Suppl 1, S4 (2010).
  15. Afgan, E. et al. *Harnessing cloud computing with Galaxy Cloud*. *Nat. Biotechnol.* 29, 972–974 (2011).
  16. Grünig, B. et al. *Practical Computational Reproducibility in the Life Sciences*. *Cell Syst* 6, 631–635 (2018).
- Batut, B. et al. *Community-Driven Data Analysis Training for Biology*. *Cell Syst* 6, 752–758.e1 (2018).





## ISCB OFFICERS AND EXECUTIVE COMMITTEE

---

Thomas Lengauer, Ph.D.  
President,  
Christine Orengo, Ph.D.  
President-elect  
Bonnie Berger, Ph.D.  
Vice President  
Yana Bromberg, Ph.D.  
Vice President  
Terry Gaasterland, Ph.D.  
Vice President  
Bruno Gaeta, Ph.D.  
Treasurer  
Scott Markel, Ph.D.  
Secretary

## ISCB BOARD OF DIRECTORS

---

Alex Bateman, Ph.D.  
Bonnie Berger, Ph.D.  
Michelle Brazas, Ph.D.  
Yana Bromberg, Ph.D.  
Catherine Brooksbank, Ph.D.  
Iddo Friedberg, Ph.D.  
Terry Gaasterland, Ph.D.  
Bruno Gaeta, Ph.D.  
Wataru Iwasaki, Ph.D.  
Janet Kelso, Ph.D.  
Thomas Lengauer, Ph.D.  
Scott Markel, Ph.D.  
Francisco Melo Ledermann, Ph.D.  
Tijana Milenkovic, Ph.D.  
Yves Moreau, Ph.D.  
Nicola Mulder, Ph.D.  
Christine Orengo, Ph.D.  
Predrag Radivojac, Ph.D.  
Farzana Rahman, Ph.D.  
Christian Schönbach, Ph.D.  
Russell Schwartz, Ph.D.  
Hagit Shatkay, Ph.D.  
Saurabh Sinha, Ph.D.  
Alfonso Valencia, Ph.D.  
Xuegong Zhang, Ph.D.

## UPCOMING AFFILIATE CONFERENCES

### MAY INSTITUTE ON COMPUTATION AND STATISTICS FOR MASS SPECTROMETRY AND PROTEOMICS

Apr 27, 2020 through May 08, 2020

<https://computationalproteomics.khoury.northeastern.edu/>

### 16TH INTERNATIONAL SYMPOSIUM ON BIOINFORMATICS RESEARCH AND APPLICATIONS

Jun 01, 2020 through Jun 04, 2020

ISCB Member Discount: 20 percent

<https://isbra.confreg.org/>

### 3RD INTERDISCIPLINARY SIGNALING WORKSHOP

Jul 20, 2020 through Jul 24, 2020

ISCB Member Discount: 15 percent

<http://signalingworkshop.org/>

### THE 9TH NATIONAL CONFERENCE ON BIOINFORMATICS AND SYSTEMS BIOLOGY OF CHINA

Oct 08, 2020 through Oct 11, 2020

ISCB Member Discount: 20 percent

<http://ccbsb2020.dtc.es.com/en.html>

[WWW.ISCB.ORG](http://WWW.ISCB.ORG)

CONNECTING, TRAINING, EMPOWERING,  
WORLDWIDE