

## VI シンポジウムの開催



## VI シンポジウムの開催

生命情報学教育の一環として、独立行政法人日本学術振興会国際事業部の支援を受け、シンポジウム「Bioinformatics beyond Omics Data Analyses – For Collaborative Effort to Expand Bioinformatics Research and Education in East Asia –」を平成 21 年 12 月 3～4 日に開催した(資料 15)。海外からは中国、韓国、台湾、ベトナム、タイ、および英国の生命情報学教育研究に携わる方々が、国内からは東北大学、九州工業大学、国立遺伝学研究所、および本学の生命情報学の教育研究に携わる研究者に講演をしていただくことができた。講演を通して、それぞれの機関の生命情報学の教育が、解析ツールのユーザを養成するのか、あるいは解析ツールの開発者を養成するのかの違いがあることが明らかになった。また、短期に人材を養成できる長所をもつ教育システム、基礎よりも実践に重きを置く教育システム、養成に時間を要する短所を抱えながらも周辺技能(英語や実験生物学)も含めた総合的な教育システムなどが、それぞれの国で行われていることが明らかになった。学内外からの参加者総計は延べ 104 名(学生 61 名、一般 43 名)の盛況な会合となり、本学内における生命情報学の関心の高さが数字であらわれることとなった。本学学生に生命情報学がかなり定着してきたことと共に、本学教員は学生のために、生命情報学を学んだ学生が国際舞台でもリーダーとして活躍できる素地を形成していかなければならない責任があることが明らかになった。このシンポジウムをきっかけとして、本学の生命情報学教育が東アジア諸国の生命情報学教育の一助になることが期待されている。

シンポジウムの概要は、以下の通りである。



シンポジウム講演者一同

## **Bioinformatics beyond Omics Data Analyses**

- For Collaborative Effort to Expand Bioinformatics Research and Education in East Asia -

**Sponsor:** Center for Informational Biology, Ochanomizu University and Japan Society for the Promotion of Science

**Co-sponsor:** Education for Women as Leaders: Computational Biology, Ochanomizu University

**Date:** 3-4, December, 2009

**Place:** Meeting Room, Science Building #3, Ochanomizu University, Tokyo, Japan

**Aim:** Bioinformatics has increased its importance in post genomic era. In the bioinformatics, integrated understanding of computation and biology is required and the education of the new generation is getting harder. To overcome the difficulty, collaboration both in education and practice is evidently required. This workshop functions as an introduction to the education systems and research activities in East Asian countries. The participants are expected to talk about both topics from their point of view. Based on the talks, we wish that the workshop functions as one step of the future network for education and researches in computational biology in East Asia.

### **Program:**

#### **3rd Dec.**

13:00 Opening by a Vice President of Ochanomizu University

13:10 Greeting from a Representative of JSPS

Chaired by Kei Yura (Ochanomizu, Japan)

13:20 Sanghyuk Lee (Ewha, Korea)

**Bioinformatics Programs at Ewha Womans University and Our Recent Work on Cancer Bioinformatics: Transcriptome Analysis for Identifying Fusion Genes**

14:00 Jun Sese (Ochanomizu, Japan)

**Biological Knowledge Discovery by Combining Large and Complex Datasets**

14:40 Hiroshi Akashi (NIG, Japan)

**Biosynthetic Constraints and *Drosophila* Genome Evolution**

15:20 Break

Chaired by Jingchu Luo (Peking, China)

15:50 Carmay Lim (Academia Sinica, Taiwan)

**Physico-Chemical Principles Governing Biological Processes**

16:30 Akinori Sarai (KIT, Japan)

**Distance Education and Research Collaboration in the Network Era**

17:10 Marasri Ruengjitchatchawalya (King Mongkut's, Thailand)

**Bioinformatics Education System and Omics Research at KMUTT**

18:00 **A Buffet-Style Welcome Party**

#### **4th Dec.**

Chaired by Sanghyuk Lee (Ewha, Korea)

10:00 Sangsoo Kim (Soongsil, Korea)

**Bioinformatics Course at Soongsil University and Our Recent Progress on Bio-Data Mining**

10:40 Kengo Kinoshita (Tohoku, Japan)

**Multi-Dimensional Correlations for Gene Coexpression and Application to the Large-Scale Data of *Arabidopsis***

11:20 Gos Micklem (Cambridge, UK)

**InterMine: A System for Large Scale Biological Data Integration**

12:00 Canteen Lunch

12:00 **Poster Session**

Chaired by Carmy Lim (Academia Sinica, Taiwan)

14:30 Oanh TP Kim (VAST, Vietnam)

**Outlook of Bioinformatics Research in Vietnam**

15:10 Jingchu Luo (Peking, China)

**ABC – Applied Bioinformatics Course for Graduate Student in Biology**

15:50 Kei Yura (Ochanomizu, Japan)

**Bioinformatics Course for Undergraduate in Ochanomizu University and Our Recent Work on Structural Bioinformatics: The Interwinding Nature of Protein-Protein Interfaces and Its Implication for Protein Complex Formation**

16:30 Closing

#### **Poster Presentation Title List**

**01: Bioinformatics Education at the National Institute of Genetics in Mishima**

Hiroshi Akashi

**02: Bioinformatics Education Course in Ochanomizu University: Graduate and Undergraduate Courses**

Kei Yura, Jun Sese

**03: A Course for Biological Knowledge Discovery by Handling Genome-Wide Data**

Satoko Kaneko, Jun Sese

**04: University Education Internationalization Promotion Program: International Cooperation on Interdisciplinary Biomedical Science Education**

Keiko Miyamoto, Kei Yura

**05: Cambridge iGEM 2009**

Gos Micklem

**06: InterMine: Open Source Data Warehouse and Query Interface**

Richard Smith, Sergio Contrino, Hilde Janssens, Jakub Kulaviak, Rachel Lyne, Kim Rutherford, Julie Sullivan, Dan Tomlinson, Matthew Wakeling, Xavier Watkins, Gos Micklem

**07: ChimerDB 2.0 – A KnowledgeBase for Fusion Genes**

Pora Kim, Suhyeon Yoon, Namshin Kim, Sanghyun Lee, Minjeong Ko, Haeseung Lee, Hyunjung Kang, Jaesang Kim, Sanghyuk Lee

**08: GARNET: Gene Set Analysis with Annotation Network Tools**

Kyoohyoung Rho, Bumjin Kim, Youngjun Jang, Sanghyun Lee, Taejeong Bae, Jihae Seo, Chaehwa Seo, Jihyun Lee, Wan Kyu Kim, Ungsik Yu, Sunghoon Kim, Sanghyuk Lee

**09: RESOPS: A Database of RNA Editing Sites in Plant Organelles Mapped onto Protein Three-Dimensional Structures**

Sintawee Sulaiman, Kei Yura, Yosuke Hatta, Masafumi Shionyu, Mitiko Go

**10: A Structural Systems Biology Approach Transcription Factor Network Inference**

Taeho Kim and Daron M. Standley

**11: Mutation@A Glance: A New Integrated Bioinformatics Tool for Analyzing Human Disease Mutations**

Atsushi Hijikata, Osamu Ohara

**12: A New Computational Method for Protein Folding Screening from NMR Spectra**

Hiromi Arai, Satoru Watanabe, Takanori Kigawa, Masayuki Yamamura

**13: A New SVM-based Algorithm for Multi-Task Learning**

Tsuyoshi Kato, Hisashi Kashima, Masashi Sugiyama, and Kiyoshi Asai

**14: Computational Identification of Plant Transcription Factors and the Construction of the PlantTFDB Database**

Kun He, An-Yuan Guo, Ge Gao, Qi-Hui Zhu, Xiao-Chuan Liu, He Zhang, Xin Chen, Xiaocheng Gu, Jingchu Luo

**15: Development of A New Biclustering Algorithm for Gene Expression Analysis by Optimizing Average Correlations with Genetic Algorithm**

Sutheeworapong Sawannee, Obayashi Takeshi, Ota Motonori, Kinoshita Kengo

**Abstracts:****Bioinformatics Programs at Ewha Womans University and Our Recent Work on Cancer Bioinformatics: Transcriptome Analysis for Identifying Fusion Genes**

*Sanghyuk Lee (Ewha Research Center for Systems Biology (ERCSB)*

*Division of Life and Pharmaceutical Sciences, Ewha Womans University, Seoul, Korea)*

**Education:**

Ewha Womans University has started an inter-disciplinary program of bioinformatics for undergraduate students as early as 2005. The curriculum consists of 34 lectures provided by the departments of biology, computer science, statistics, mathematics, and chemistry. Students may select the degree as one of dual majors or minor according to the credits. Even though we do not have a separate degree program for graduate level studies, all related professors belong to the Ewha Research Center for Systems Biology (ERCSB) that was established in 2008. ERCSB's research is focused on cancer systems biology with strong emphasis on systems bioinformatics.

**Research:**

Developing diagnostic and prognostic biomarkers is an important field in cancer research. Typically, omics experiments are carried out to identify candidate genes, transcripts, or proteins. In this presentation, I will describe our recent activities to identify fusion genes that play important roles in tumorigenesis and cancer progress. We recently developed a computational pipeline to identify fusion transcript candidates by analyzing genomic alignment of transcripts. Analysis of transcriptome data in the GenBank and SRA (short read archives) produced ample candidates. We also collected the literature data (journal publications) on fusion genes. Fusion candidates from computational pipeline and literature cases were all integrated into a database called ChimerDB 2.0, available at <http://ercsb.ewha.ac.kr/fusiongene>. In addition, we performed a transcriptome sequencing of lung cancer cell lines using the Roche 454 GS FLX technology and identified several novel candidates. Another major part of my research is building an integrated knowledgebase that encompasses genome, phenome, and perturbagenomes. I will briefly describe the contents of integrated knowledge and show an application example.

**Reference:**

Pora Kim; Suhyeon Yoon; Namshin Kim; Sanghyun Lee; Minjeong Ko; Haeseung Lee; Hyunjung Kang; Jaesang Kim; **Sanghyuk Lee** (2009) ChimerDB 2.0 – a knowledgebase for fusion genes updated. *Nucleic Acids Research*, in press (doi: 10.1093/nar/gkp982).

## **Biological Knowledge Discovery by Combining Large and Complex Datasets**

*Jun Sese(Ochanomizu University, Japan)*

### **Education:**

Ochanomizu Univ. has four years history about bioinformatics education programs for graduate students (<http://bioinfo.is.ocha.ac.jp/>, in Japanese). The programs include cutting-edge bioinformatics topics, basic programming skills and usage of bioinformatics tools. More than 50 students whose majors are biology, chemistry and computer sciences finished this course. To extend this course to the international education, we have recently started a program about international cooperation on interdisciplinary biomedical science education that is collaboration between Tokyo medical and dental University and Ochanomizu University. In this program, we provide classes about computational biology by foreign top researchers (<http://cib.cf.ocha.ac.jp/KASOKU/index-e.html>).

### **Research:**

Recent biotechnologies such as next generation sequencers give us huge and complex dataset. Aims of our laboratory are the development of novel computational methods to analyze such datasets and the discovery of new biological and medical knowledge by application of the methods. Our recent research topics are (1) finding active protein-protein interaction (PPI) network in specific conditions, (2) development of search engine for gene expression data and (3) analysis of carbohydrate-macromolecule interactions observed by glycan arrays that are devices for high-throughput screening of the interactions. The common method for the analyses is to handle simultaneous analysis of graph structures and their associated features. For example, in PPI network analysis, we combined PPI network with yeast gene expressions under various stress environments observed by microarrays, and discovered sub-networks whose genes are highly activated under specific stresses. By using this method, we identified a heat-shock related sub-network and two independent networks highly expressed under oxidative stress environments, which are identical to proteasome complexes 19S and 20S.

## **Biosynthetic Constraints and *Drosophila* Genome Evolution**

*Hiroshi Akashi(Department of Population Genetics,  
National Institute of Genetics, Mishima Japan)*

### **Education:**

The National Institute of Genetics welcomes graduate students as part of the Sokendai Graduate University of Advances Studies. Bioinformatics research and education is central to our institute's mission and the DNA Data Bank of Japan (DDBJ) is housed on



campus. Interaction between empirical and computational researchers is particularly strong at NIG and I will present some examples. In addition, “internationalization” is strongly emphasized and all graduate courses and almost all institute seminars are conducted in English.

#### **Research:**

Mutations with subtle effects on cell physiology may play an important role in biological adaptation and genome evolution. However, the functional effects of nearly neutral mutations may lie outside the detection limits of experimental systems. I will present evidence that a considerable fraction of the *Drosophila* genome is involved in adaptation for efficient biosynthesis. Translational optimization is well documented for synonymous codon usage. In most genomes, codon usage is biased toward "major" codons for each amino acid. Abundant tRNAs recognize major codons; translation at such codons is faster and more accurate than at minor codons. Although the physiological impact of synonymous changes may be vanishingly small, tRNA pools and codon usage are co-adapted to allow efficient use of ribosomes and to limit the costly production of dysfunctional peptides. Both relationships between codon bias and gene expression and population genetic analyses support "major codon preference" in multi-cellular eukaryotes. However, adaptation for efficient biosynthesis at other classes of DNA sites has received little attention. Associations between gene expression and amino acid composition as well as intron size suggest that many sequence features may function to enhance biosynthesis. Genome-scale analyses of closely related *Drosophila* lineages show that mutation patterns and selection intensity for codon bias vary frequently on the time-scale of molecular evolution. Strong departures from steady-state amino acid composition in lineages showing changes in codon usage suggest that similar forces may affect silent and protein evolution.

#### **Physico–Chemical Principles Governing Biological Processes**

*Carmay Lim(Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan)*

Our research interests are to

- (i) unravel the underlying physico–chemical principles governing biological processes,
- (ii) develop new methods, as required, along the way and
- (iii) use the above principles and methods to design molecules with potential therapeutic utility.

This talk will provide an example of our work in these three areas.

## References:

1. Metal Binding Affinity and Selectivity in Metalloproteins: Insights from Computational Studies. Todor Dudev & Carmay Lim\*, *Annual Reviews in Biophysics*. (2008) 37:97–116
2. Physical Basis of Structural and Catalytic Zn-binding Sites in Proteins. Yu-Ming Lee & Carmay Lim\*, *J. Mol. Biol.* (2008) 379:545–553.
3. Predicting DNA-Binding Amino Acid Residues from Electrostatic Stabilization upon Mutation to Asp/Glu and Evolutionary Conservation. Yao Chi Chen, Chih Yuan Wu & Carmay Lim\*, *Proteins: Structure, Function & Bioinformatics* (2007) 67: 671–680.
4. Predicting RNA-binding sites on proteins based on electrostatics, evolution, and geometry. Yao Chi Chen & Carmay Lim\*, *Nucleic Acids Res.* (2008) 36(5): e29.
5. Common Physical Basis of Macromolecule-Binding Sites in Proteins. Yao Chi Chen & Carmay Lim\*, *Nucleic Acids Res.* (2008) 36:7078-7087.

## Distance Education and Research Collaboration in the Network Era

Akinori Sarai(Kyushu Institute of Technology, Japan)

Explosive amount of data have been produced by genome-scale analysis of biological system. However, because of the speed of data increase, the extraction of knowledge from the data is becoming a rate-limiting step, and the data are not efficiently utilized in research. Bioinformatics has emerged to circumvent this problem, and is expected to make a huge contribution not only to the advancement of basic biological science but also to biotechnology and medicine in the future. However, because of its rapid growth, the developments of infrastructure and education fell behind. Bioinformatics is a highly interdisciplinary field of science, covering biology, information science, mathematics, physics, chemistry and engineering. Thus, this is the field that needs interaction of people with different academic backgrounds and cooperation among them. However, the lack of supporting systems to facilitate such cooperation is one of the factors to hinder the sound development of this field.

Asia is expected to grow in economy in near future. However, although Asia holds more than half of the world population, many of the countries are in the development stage. In particular, education and research in rapidly growing fields of science such as bioinformatics are short of infrastructures (facilities, professional people, etc.) and financial source. Students and researchers in remote area have difficulties in travel to attend conferences and workshops. On the other hand, the impressive development of Internet has made people to share information over geographical barriers. So, this would change the way of education and research in future. Traditional education is restricted in

space and time; i.e., recruiting teachers and students at one place at specific time. On the other hand, the distance education can reach out anyone connected by the Internet.

Research collaboration would also be accelerated by sharing scientific data and idea through the Internet.

We have started to hold virtual on-line training workshop on bioinformatics several years ago. The workshops had been open for two months each year, and registered participants can access to the local streaming servers to view the multimedia lectures contributed by experts in Asian countries, and they can also ask teachers questions and get answers any time during the period. In the last workshop, more than 1,000 people participated from 51 countries. We are also developing a system where people can exchange idea and opinion on various issues of bioinformatics in order to facilitate research collaboration.

In this talk I will discuss the distance education and research collaboration. I will also talk about some of my current research in bioinformatics.

### **Bioinformatics Education System and Omics Research at KMUTT**

*Marasri Ruengjitchatchawalya<sup>1,2</sup>, Apiradee Hongsthong<sup>3</sup>, Asawin Meechai<sup>1,4</sup> and  
Supapon Cheevadhanarak<sup>1,2</sup>*

*(<sup>1</sup>Bioinformatics and Systems Biology Program, KMUTT, Bangkok 10140, Thailand*

*<sup>2</sup>School of Bioresources and Technology, KMUTT*

*<sup>3</sup>Biochemical Engineering and Pilot Plant Research and Development Unit (BEC)*

*Cooperative unit between BIOTEC and KMUTT*

*<sup>4</sup>Department of Chemical Engineering, Faculty of Engineering, KMUTT)*

Bioinformatics program (BIF) at King Mongkut's University of Technology Thonburi (KMUTT), a cooperative international Master's degree program by School of Bioresources and Technology (SBT) and School of Information Technology (SIT), has been launched in 2003. This program has been designed for students who desire focused training in the elements of computer science, biology and biochemistry needed for a successful career in this exciting multidiscipline. Students in our program received comprehensive training in genomics, algorithms for sequence analysis, database design and management, software engineering and programming (including web-based development). Each student will apply their skills to a practical project, where they designed and implemented a solution to a real-world problem under the guidance of experienced mentor(s) in academia and/or industry. In 2004, KMUTT in collaborated with Thai and international universities/research institutes, and Grant-in-aid in part supported by the National Center for Genetic Engineering and Biotechnology (BIOTEC), The National Science and Technology Development Agency (NSTDA) for 5 years has

been operated the program as a scholarship program. The program enable students to have an internship for 5-6 months at leading universities/ research institutes, such as University of Washington, USA; Denmark Technical University, Denmark; The Hospital for Sick Children, Program in Genetics & Genomic Biology, Canada; IMB, University of Queensland, Australia; Osaka University, Japan; Beijing Institute of Genomics, China; National University of Singapore, Genome Institute of Singapore, Singapore; Bioinformatics and Data Management for Research Unit, Siriraj Hospital, Mahidol University, Thailand; and Bioinformatics Laboratory, BIOTEC, Thailand. The recent status of BIF's alumni is mostly working in Bioinformatics research units in Thailand and international universities/ research institutes.

Bioinformatics research, including Systems Biology and Omics research, at KMUTT are the forerunner in Thailand. The research groups have focused on systems analyses to analyze metabolic networks of various organisms such as *Spirulina platensis*, *Saccharomyces cerevisiae*, *Cassava*, *Plasmodium*, and *Tuberculosis*, aiming to understand biology of these organisms at the system level with applications to strains improvement, high value chemicals production and drug targets identification. Moreover, proteomic analyses in response to stress conditions including temperature stresses were performed in *S. platensis*. The results obtained from these studies in coupled with *Spirulina* genome sequence database lead to the completion of *Spirulina* database, which will be very beneficial to the science community especially in the field of cyanobacteria and plants. Furthermore, the information available in genome and proteome databases could make the study of protein-protein interaction network of *Spirulina* possible. Therefore, the construction of protein-protein interaction network of *Spirulina* under temperature stress condition is now in progress.

資料15 「Bioinformatics beyond Omics Data Analyses – For Collaborative Effort to Expand Bioinformatics Research and Education in East Asia –」開催案内

ワークショップ  
タンパク質の発現・相互作用解析を越えたバイオインフォマティクスの展開  
ー東アジアにおける生命情報学の研究と教育の協力をめざしてー

# Bioinformatics beyond Omics Data Analyses

For Collaborative Effort to Expand Bioinformatics Research and Education in East Asia

**Date: 13:00, 3rd - 17:00, 4th, Dec., 2009**  
**Place: Ochanomizu University**

Latest information at <http://clb.cf.ocha.ac.jp/symposium-e.html>

Invited Speakers (alphabetical order):  
Hiroshi Akashi (NIG)  
Oanh TP Kim (VAST)  
Sangsoo Kim (Soongsil)  
Kengo Kinoshita (Tohoku)  
Sanghyuk Lee (Ewha)  
Carmay Lim (Academia Sinica)  
Jingchu Luo (Peking)  
Gos Micklem (Cambridge)  
Marasri Ruengjitchachawalya (King Mongkut's)  
Akinori Sarai (KIT)  
Jun Sese (Ochanomizu)  
Kei Yura (Ochanomizu)

**Toward integrated study and integrated education**

**Presentation of recent research and introduction of education system for computational biology in East Asian countries**

Poster presentation is welcomed.  
Please contact the organizer.

Sponsor: Center for Informational Biology, Ochanomizu University  
Japan Society for the Promotion of Science  
Co-sponsor: Education for Women as Leaders Program: Computational Biology, Ochanomizu University

主催: お茶の水女子大学 生命情報学教育研究センター  
日本学術振興会  
共催: お茶の水女子大学 女性リーダー育成プログラム 「生命情報学をつかいこなせる女性人材の育成」

Contact: Kei Yura ([yura.kei@ocha.ac.jp](mailto:yura.kei@ocha.ac.jp))

ワークショップ  
タンパク質の発現・相互作用解析を越えたバイオインフォマティクスの展開  
ー東アジアにおける生命情報学の研究と教育の協力をめざしてー

# Bioinformatics beyond Omics Data Analyses

For Collaborative Effort to Expand Bioinformatics Research and Education in East Asia

**Date: 13:00, 3rd - 17:00, 4th, Dec., 2009**  
**Place: Ochanomizu University**

Latest information at <http://clb.cf.ocha.ac.jp/symposium-e.html>

Invited Speakers (alphabetical order):  
Hiroshi Akashi (NIG)  
Oanh TP Kim (VAST)  
Sangsoo Kim (Soongsil)  
Kengo Kinoshita (Tohoku)  
Sanghyuk Lee (Ewha)  
Carmay Lim (Academia Sinica)  
Jingchu Luo (Peking)  
Gos Micklem (Cambridge)  
Marasri Ruengjitchachawalya (King Mongkut's)  
Akinori Sarai (KIT)  
Jun Sese (Ochanomizu)  
Kei Yura (Ochanomizu)

**Toward integrated study and integrated education**

**Presentation of recent research and introduction of education system for computational biology in East Asian countries**

Poster presentation is welcomed.  
Please contact the organizer.

Sponsor: Center for Informational Biology, Ochanomizu University  
Japan Society for the Promotion of Science  
Co-sponsor: Education for Women as Leaders Program: Computational Biology, Ochanomizu University

主催: お茶の水女子大学 生命情報学教育研究センター  
日本学術振興会  
共催: お茶の水女子大学 女性リーダー育成プログラム 「生命情報学をつかいこなせる女性人材の育成」

Contact: Kei Yura ([yura.kei@ocha.ac.jp](mailto:yura.kei@ocha.ac.jp))

