Strategic Programs for Innovative Research Field 1 Supercomputational Life Science

# BioSupercomputing Newsletter

2014.3 Vol.10


## CONTENTS

# Toward integrated simulation of biological systems beyond hierarchy

## Developing models for reproducing and predicting biological phenomena of all scales, from molecular motions to behaviors of tissues and organs

Theme 3 "Hierarchical integrated simulation for predictive medicine" (leader: Prof. Shu Takagi, The University of Tokyo) aims to integrate simulators of biological systems (thrombosis, heart, muscles and skeleton, cerebral nervous system, etc.) that were separately developed in the grand challenge program that ended last spring in order to reproduce the complicated processes involved in various diseases including myocardial infarction, and contribute to the prediction of pathology and treatment of these diseases. The new integrated simulation model to be developed is required to achieve simultaneous reproduction on both the multi-scale (reproduction of biological phenomena of all scales from molecular level to the entire body) and multi-physical (including mechanical, biochemical and electrical phenomena) levels by making effective use of the powerful computer resources of the K computer. We interviewed three researchers who are working toward the development of a simulation system for Theme 3 while confronting many issues.

Research Associate, School of Engineering, The University of Tokyo — Kazuya Shimizu
Research Associate, School of Engineering, The University of Tokyo — Naoto Yamamura
Research Associate Professor, Graduate School of Frontier Science, The University of Tokyo — Takumi Washio

## Models to be constructed by hierarchical integration

**In what actual research are you engaged?**

Shimizu: Dr. Yamamura and I are developing an integrated nervous and musculoskeletal simulation model of the entire human body, which can reproduce micro- to macro-scale activities that occur in the body, aiming to elucidate motor dysfunctions caused by neurological disorders, such as Parkinson's disease. The motor command from the brain is modified in various ways in the signaling pathway, such as the spinal cord, and finally induces motion, such as contraction of a muscle and articular movement. One of our goals is to reproduce such a series of processes with a computer, and contribute to unveiling the causes and treatment of diseases. I am mainly in charge of constructing a neural network model in the spinal cord.

Yamamura: I am developing a finite element simulator of the musculoskeletal system. The musculoskeletal simulator receives motor commands from the spinal cord and generates a joint movement as a result of muscle contraction. Our collaborators are developing simulators of the central nervous system for generating the motor commands. We will finally integrate simulators of the brain, spinal cord and musculoskeletal system.

Washio: Our group is developing a heart simulator (UT-Heart), which can reproduce a living heart by using the K computer. At present, we are modeling cardiac beat. Basically, we are trying to reproduce the pulsation of the heart by solving equations that follow the laws of physics based on the movements of each molecule. However, individual molecules do not move alone, but affect each other and are also affected by peripheral muscles. Thus, it is difficult to find equations that describe average movements of the molecules. Besides, we need to not only reproduce the movements of the myocardia by calculating the forces acting on the molecules and larger macroscopic entities, but also to solve the interactions between the blood, which is fluid pumped out from the heart, and the ventricular wall. The flow of blood cannot be calculated only for the heart alone - the flow in the entire body needs to be reproduced. Therefore, the huge capacity of the K computer is required to achieve consistent computation from molecular motions to systemic circulation, otherwise we cannot reproduce the phenomena well.

Yamamura: We are also developing models for skeletal muscle contraction from the molecular level, like the myocardial model which Dr. Washio developed.

Shimizu: Neurons exchange electrical signals by closing and opening channels on the membranes of cells, and changing ion concentrations. Thus, for signal transmission simulation, we are also constructing models which may not be as small as molecules, but are much smaller than cells. Compared to the simulation of the brain, we do not need large computer resources for calculating the necessary signaling for moving a joint. However, when we think about reproducing movements that involve the entire body, such as "walking and running", even the spinal cord model will need a considerably large computer resources.

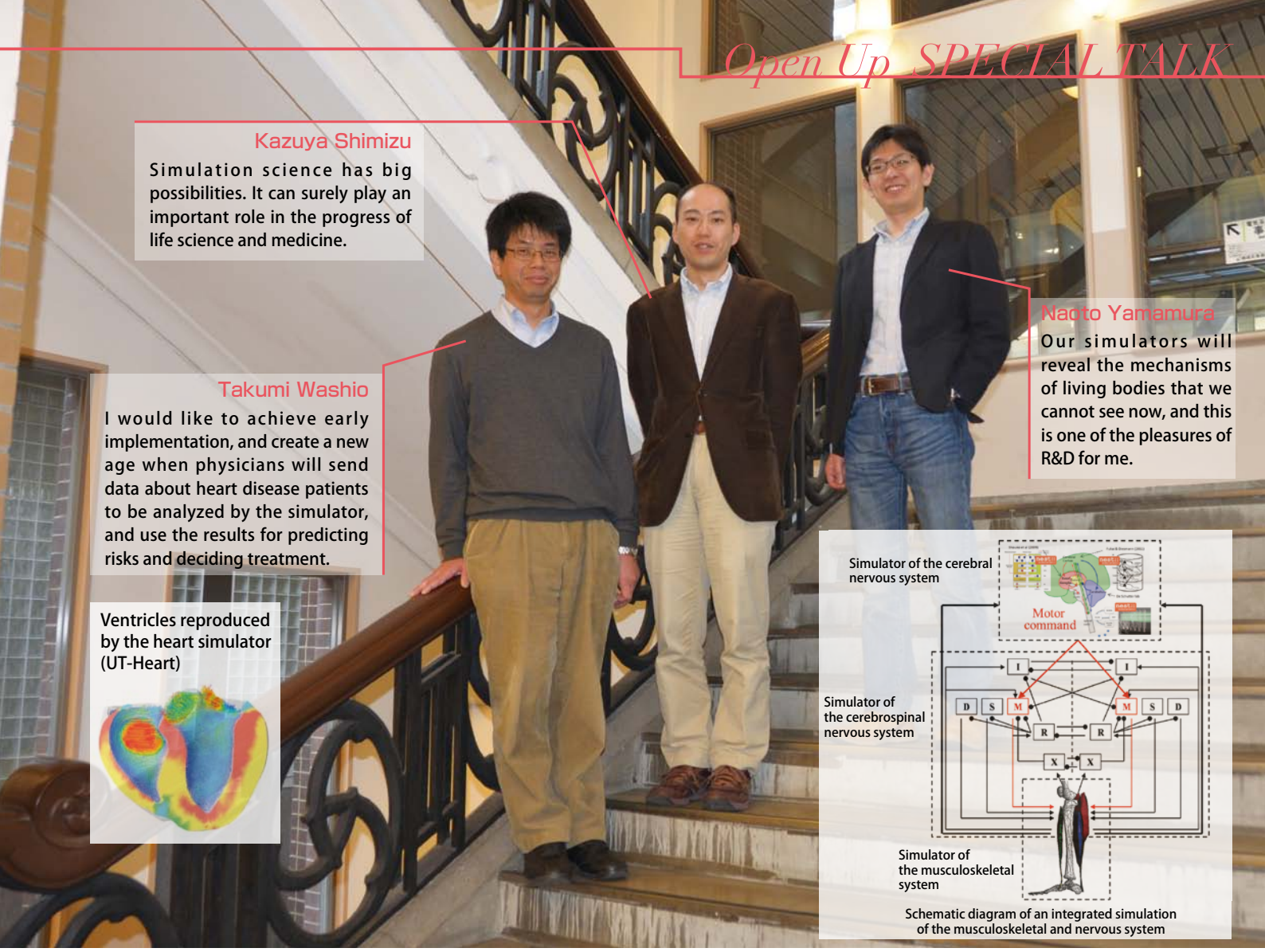
## Basic information for modeling is difficult to acquire

**What are the most difficult points for constructing the simulators?**

Shimizu: Before getting involved in life science, I mainly studied simulation of fluids. For example, in engineering simulation, basic and detailed blueprints are available, and it is relatively easy to perform material tests for collecting information required for modeling, such as resistance and thermal conductivity. But in the case of the human body, we cannot open the body and measure the values; and thus it is difficult to obtain basic information.

Yamamura: I had developed a simulator for sheet metal forming process before I started to study living bodies, and I also know the difficulty of collecting basic data on the human body. We particularly want to measure the "living human body", and want to use non-invasive measurement methods, but it is difficult. The differences among individuals are also large. We are using MRI data for the shapes of skeletal muscles, but MRI does not show the thin aponeurosis, which transmits the force of muscle contraction to the bone. We have to consult anatomical references for data that we cannot measure.

Shimizu: I think the nervous system is even more difficult to measure.



**Kazuya Shimizu**

Simulation science has big possibilities. It can surely play an important role in the progress of life science and medicine.

**Takumi Washio**

I would like to achieve early implementation, and create a new age when physicians will send data about heart disease patients to be analyzed by the simulator, and use the results for predicting risks and deciding treatment.

**Naoto Yamamura**

Our simulators will reveal the mechanisms of living bodies that we cannot see now, and this is one of the pleasures of R&D for me.

Ventricles reproduced by the heart simulator (UT-Heart)






Schematic diagram of an integrated simulation of the musculoskeletal and nervous system

**Washio:** The difficulty is the same for the heart. For example, CT and MRI data are available for the shape of the heart, but the heart is always beating. We do not know the accurate natural shape of the living human heart, and it varies among individuals. Myocardia also have a very complicated structure with myofibers running through them while making twists. The heart has a very clever structure, and ejects about 60% of the blood just by 15% contraction of the myocardia. In order to know its detailed structure with current measurement techniques, we need to dissect the body or stop the heart. What we can do now is to collect as much data and experimental results as possible, and construct a model that can reproduce all characteristic phenomena.

## Toward development of simulators that can contribute to society

**Washio:** Luckily, the K computer has improved the reliability of heart simulation. It has aroused interest among experimental molecular biologists and cardiac surgeons in simulation. Cooperative relations with them are expanding, and they are helping us in experiments and providing data. There have also been inquiries about the availability of simulation for deciding an operational method for treating patients with an innate heart defect. A challenge has already started to unveil the mechanism of developing hypertrophic cardiomyopathy by using the heart simulator. I believe that the heart simulator will be useful for both clinical and basic medicine, and the results will further improve the simulation models.

**Yamamura:** It is wonderful that simulation scientists are cooperating with medical personnel and molecular biologists, and enhance each other. The integrated simulator for the nerves and entire musculoskeletal system is still in the stage of developing models, but I believe it will attract the attention of medical personnel and scientists.

**Shimizu:** For example, the mechanism by which Parkinson's disease develops is not fully understood. The situation will be greatly altered if simulators reveal the mechanism and provide useful information for deciding treatment plans. It is one of our important roles to show that computer simulation is useful for development of new medicines and new medical treatments.

**Washio:** Hierarchical integrated simulation may connect physicians in medical settings and molecular biologists, between whom there is little communication at present. For this, it is important to improve the reliability of simulators.

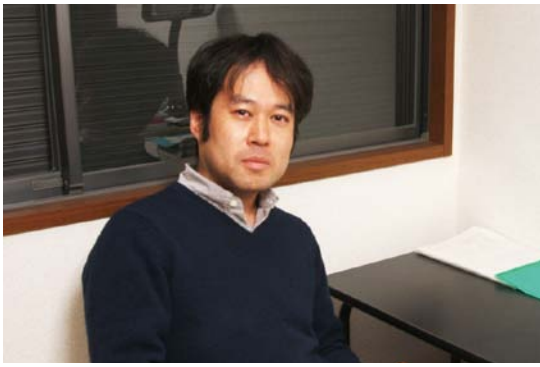
**Yamamura:** So we have to produce reliable results that can contribute to society.

**Washio:** I think it is the goal for us, who are involved in this project, to develop and elaborate simulators that are not mere tools for writing research papers, but are acknowledged by people as a useful technology.

Theme 1 Simulations of biomolecules under cellular environments

# Exploring the biomolecular world through molecular dynamics free energy calculation

Quantum Beam Science Directorate, Japan Atomic Energy Agency
Yoshiteru Yonetani



Free energy is an important quantity for understanding various phenomena of condensed molecular systems. For example, liquid water is transformed into ice as temperature decreases. This is because the free energy in the ice state becomes lower than that in the liquid sate. Free energy calculation using molecular dynamics simulation* allows us to predict such molecular phenomena.

The biomolecular world in cells, including DNA and proteins, follows the same physical rule. Thus, by calculating the free energy, we can understand various biological phenomena such as molecular recognition, structure formation and molecular transportation, However, in biological systems, free energy is difficult to calculate. This is because the molecular structure is very complex and is changing with a large fluctuation. It takes quite a long computing time to reproduce such various molecular structures.

Thanks to the recent increase of computing power such as the K computer, biomolecules became an attainable target of the free energy calculation, and many researchers are tackling related problems. We are particularly interested in the interaction between DNA and proteins. By calculating the free energy, we will reveal the molecular behavior and biological functions. Fig. 1 shows our recent calculation on a protein, lac repressor. Using the molecular dynamics method called Adaptive Biasing Force [1], we revealed not only the dissociation process of the DNA-protein complex, but also the accompanied free energy changes [2].

Our result is significant in the following two points. One is that we gained an understanding of the question, "How do DNA-binding proteins move around DNA?" Our free energy result found that dissociation of the protein from DNA is rare, and that the protein is likely to slide on the DNA. This finding concerns an unsolved problem, "How do DNA-binding proteins find their target sequence?" This problem is one of the most attractive topics in biology, and many researchers are puzzling over the problem.

The other significant point of our study is about the DNA sequence recognition. The sequence of DNA consists of four kinds of bases, A, T, G and C. Among them, the lac repressor protein binds specifically to the region with GTGAGCG. This has been clarified experimentally. However, the reason remained unclear. What is the origin for the sequence-dependent binding? Our free energy results answered this question by providing an atomic-level picture of DNA-protein binding. The result is considered to be useful for elucidating various intracellular events such as gene expression.

There remain several problems to be solved. Present computer simulation can treat only a small part of the protein and DNA, but next-generation Exa-scale supercomputing will enable us to treating much larger molecules. For example, the simulation of a compactly folded DNA structure, nucleosome, will become possible. Exa-scale supercomputing will reveal the atomic-level behavior of such larger molecules, and provide us with a more realistic picture of the DNA-protein interactions.

* Molecular dynamics simulation: Computer simulation method to derive trajectories of many-particle systems by solving Newton's equations of motion.

**References**

[1] E. Darve, D. Rodríguez-Gómez, A. Pohorille, *J. Chem. Phys.*, **128**, 144120 (2008).
[2] Y. Yonetani, H. Kono, *J. Phys. Chem. B,* **117**, 7535 (2013).

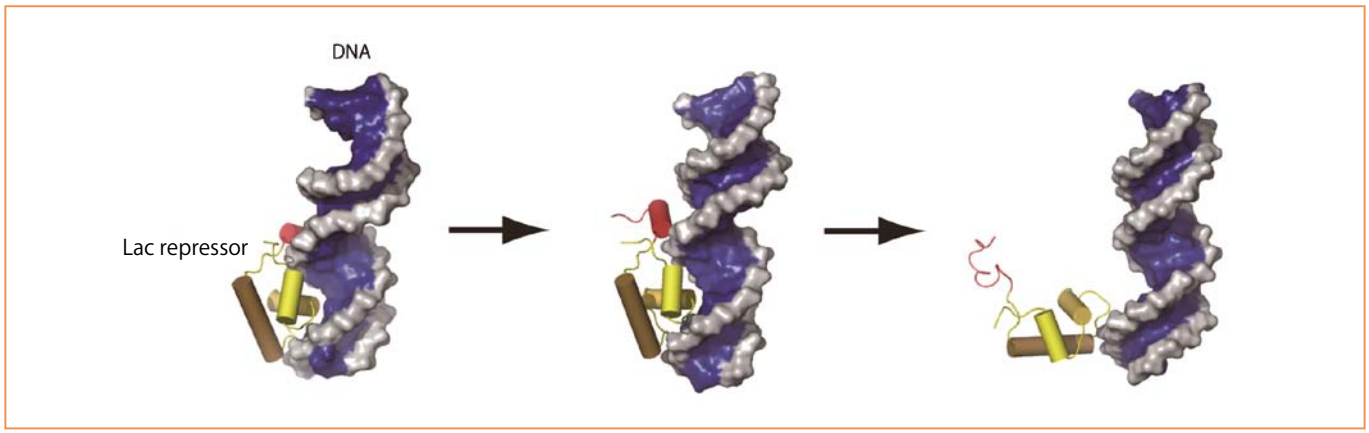



Fig. 1: Dissociation process of the DNA-lac repressor complex

See the URL on the right for detailed information on this report. http://www.kobe.riken.jp/stpr1-life/en/newsletter/Vol.10/zoomin01.html

## Theme 3 Hierarchical integrated simulation for predictive medicine

# Integrated simulation of whole-body musculoskeletal-nervous system

Department of Mechano-Informatics,
Graduate School of Information Science and Technology, The University of Tokyo
Ko Ayusawa

One of our major goals is to understand Parkinson's disease through numerical computing using a supercomputer. Parkinson's disease causes a wide spectrum of disorders including tremor in a limb, muscle rigidity and difficulty in maintenance of postural balance. Insufficiency of a neurotransmitter called dopamine in the brain is considered to be the cause. However, the mechanism responsible for systemic disturbance of motor function by dopamine insufficiency in the brain is still poorly understood. This study aims to elucidate the pathway from intracerebral phenomena to whole-body motion, and tries to clarify the mechanism of neurological disorders responsible for motor dysfunction through a large-scale simulation integrating the brain, spinal cord and whole-body musculoskeletal system. A supercomputer is necessary for such a large-scale simulation.

In whole-body integrated simulation, various simulations including the cerebrospinal nervous system, cellular-level muscle contraction and whole-body musculoskeletal motion are integrated. Therefore, several research groups with different research backgrounds are cooperating on the project. Our group has actually been working on humanoid robots. To study a humanoid robot is to understand human beings. Through modeling of the human skeleton, muscles and nerves based on robot technologies, it is possible to estimate muscular tension and identify neural reflex networks from human motion measurements. Our group has developed a whole-body musculoskeletal model with 53 bone links and 1,206 segmentalized muscles and tendons, which makes it possible to estimate muscular tension during exercise on a real-time basis (Fig. 1).

In regard to the aforementioned musculoskeletal model, we are trying to connect a detailed muscle model to a cellular-level muscle activity simulation and a cerebrospinal nerve system simulation in cooperation with other research groups. Since the aforementioned model assumes usual computational calculation, a simple wire model is used for muscles. In this research, we construct a spinal nerve system network throughout the body, and connect it to a skeletal muscle model consisting of muscle fibers. Simulation using an upper arm model which combines the simple muscle model and the spinal nerve system network can reproduce a phenomenon called the myotatic reflex on the muscles of the upper arm (Fig. 2). This phenomenon is a reflexive muscle contraction caused by muscle stretch. When a neural signal for muscle stretch reaches the spinal cord, another signal to contract the muscle is sent via the spinal cord. Only a simulation through connection of the musculoskeletal model to the spinal nerve network could reproduce such a phenomenon. By connecting it further with a brain model, we plan to realize a simulation in which brain signals activate skeletal muscles throughout the body.

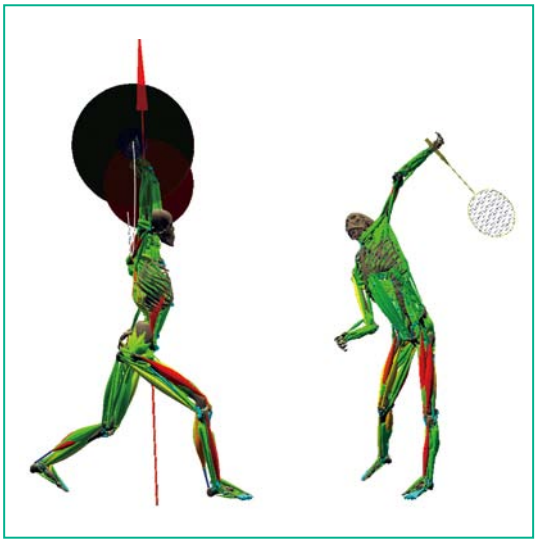


Fig. 1: Example of visual representation of muscle tension estimation. Red-colored muscles are more active than green ones.

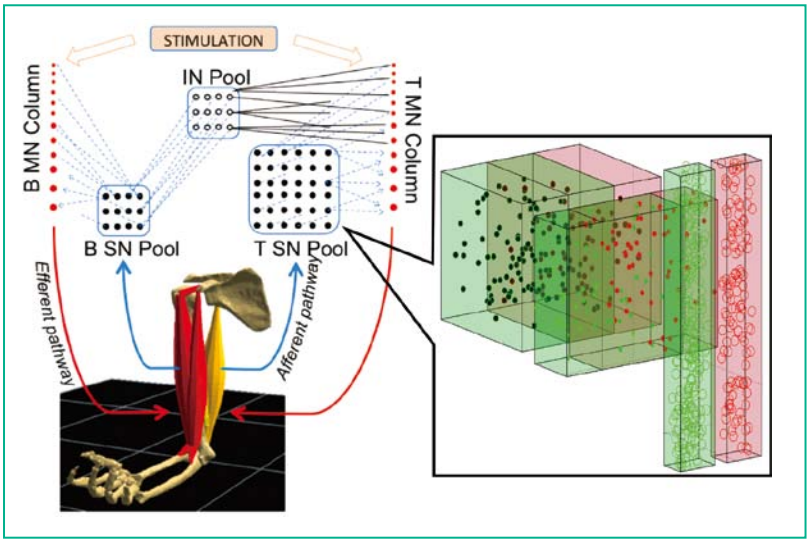


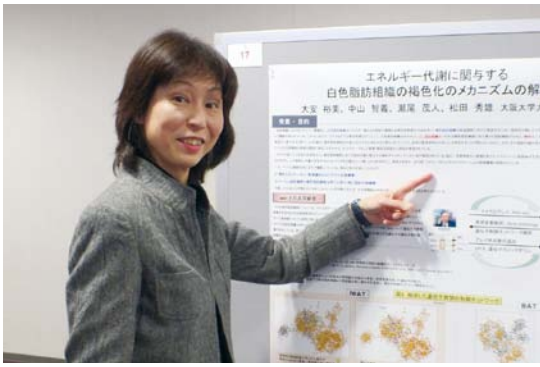

Fig. 2: Simulation connecting a musculoskeletal model of the upper arm and a spinal nerve model.

See the URL on the right for detailed information on this report. http://www.kobe.riken.jp/stpr1-life/en/newsletter/Vol.10/zoomin02.html

Theme 4 Large-scale analysis of life data

# Towards elucidation and control of the functions of brown adipocytes which contribute to energy expenditure

Department of Bioinformatic Engineering,
Graduate School of Information Science and Technology, Osaka University
Hiromi Daiyasu



Adipocytes, usually regarded as a nuisance, are roughly divided into two types, white adipocytes and brown adipocytes. White adipocytes play a role in accumulation of fat which acts as an energy source, as well as in secretion of bioactive substances. Brown adipocytes are known to have the function of consuming energy through lipolysis for heat production to regulate body temperature. Human brown adipocytes were believed to disappear after birth, but it has been revealed recently that they exist in adults, and their metabolism is closely involved in common diseases including obesity and diabetes. Therefore, they are being intensively investigated from the aspects of both basic and clinical sciences. Previous studies using mice have uncovered the parts of the genes and mechanisms involved in lipolysis by brown adipocytes. The sympathetic nervous system stimulates $\beta$-adrenergic receptors to activate its signaling pathway, which results in heat production in mitochondria by a membrane protein called UCP-1. In addition, a third adipocyte similar to the brown adipocyte has been identified, which is differentiated from the white adipocyte by cold stimulation (browning). It is called a beige adipocyte, and contributes to heat production together with the brown adipocyte (Fig. 1). However, both the browning mechanism and the relationship between brown adipocytes and beige adipocytes are still unknown. It is thought that there are several pathways leading to lipolysis, but both the master gene that dominates the whole system of energy metabolism and the overall mechanisms which regulate the system remain to be clarified. The elucidation of these mechanisms would enable control of the system, which would lead to the prevention of common diseases and improvement of symptoms.

It is known that human brown adipocytes share several characteristics with mouse beige adipocytes. Our group elucidates the function of brown adipocytes and regulatory mechanisms of energy metabolism in collaboration with Prof. Teruo Kawada and his colleagues of the Graduate School of Agriculture, Kyoto University. Kawada's lab is investigating adipocytes using mice, and measured gene expression from the adipose tissues of mice reared under cold stimulation (4ºC) by microarray and RNA-seq. Our group analyzed these data to identify the genes and their regulatory relationship, which characterize white, brown and beige adipocytes. In order to analyze expression data measured every few hours on about 25,000 genes, enormous computational resources are required, and the processing of such data is beyond the powers of conventional computers. The K computer can process this massive data, and enables estimation of the gene regulatory network (Fig. 2).

Our analyses with the K computers suggest that the genes and the pathways at work in beige adipocytes, which are different from those in brown adipocytes, rapidly induce UCP-1 to regulate fat metabolism.

We aim to obtain knowledge useful for the remedy of common diseases by elucidating the regulatory relationships of genes functioning in adipocytes with the approach described above.

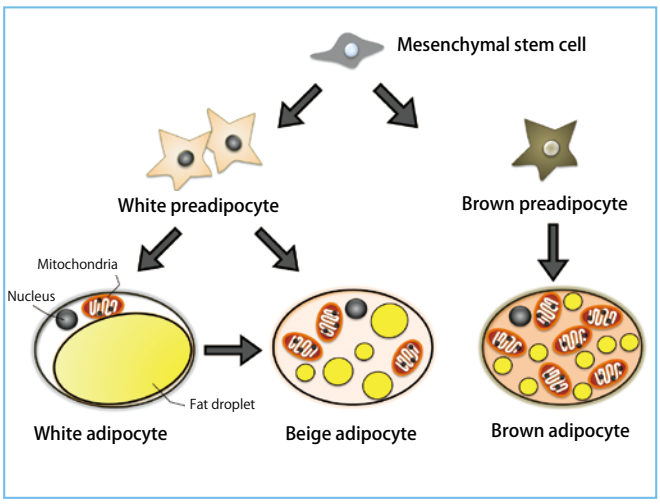



Fig. 1: Classes of adipocytes and their differentiation
Dots in mitochondria show UCP-1.

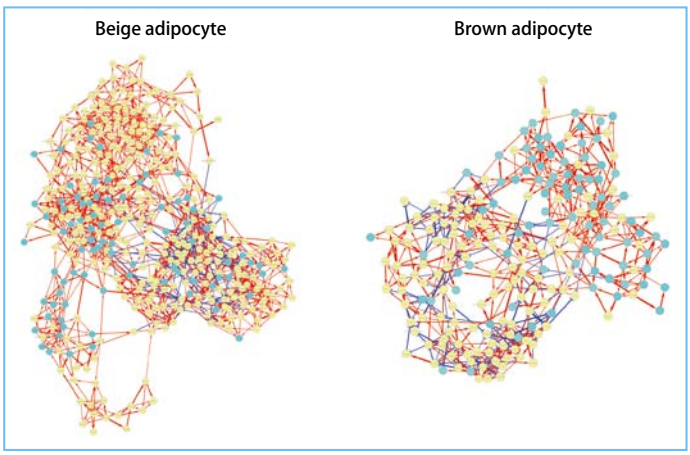



Fig. 2: Regulatory network of two types of adipocytes under cold stimulation
Genes with large differential expression are shown by circles. The red line and blue line show relationships for up-regulation and down-regulation, respectively.
Blue circles show a gene detected in both networks.

See the URL on the right for detailed information on this report. http://www.kobe.riken.jp/stpr1-life/en/newsletter/Vol.10/zoomin03.html

SCLS Gotcha!

# HPCI Human Resources Fostering Programs by the AIST Computational Biology Research Center

The Computational Biology Research Center (CBRC) of the National Institute of Advanced Industrial Science and Technology (AIST) conducts SCLS's human resources fostering programs in Odaiba in Tokyo, an artificial island which has an appearance very similar to the Port Island in Kobe. The results of research using large-scale computer systems are often reported in various meetings held in many parts of Japan. However, it is still not easy for researchers in the field of life science to use large-scale computing systems even though they wish to try them out. Therefore, CBRC decided to promote human resources development with the aim of increasing the number of users of such computing systems, and fostering people who can carry out research using large-scale computing systems like the K computer.

HPCI tutorial

The main feature of human resources fostering programs conducted here is a well-developed educational system that offers four learning styles: "workshops" that transmit the outcome and prospects of research about large-scale computing systems to the general public; "seminars" given by many leading-edge researchers; "tutorials" that provide hands-on training using one personal computer per person; and "e-learning courses" which participants can take in their spare time.

In particular, seminars are not only open to the public but also designated as a subject for which a credit may be awarded in cooperation with the Graduate School of Frontier Sciences, the University of Tokyo. They are distributed to three campuses of the University of Tokyo via video-conferencing systems. In the hope of inspiring students who listen to the same lecture as professional researchers do, including company employees, we hold these seminars every year.


Example of an e-learning screen: Slides are changed according to the progress of a lecture video.

The curriculum of tutorials is focused on offering practical exercises. It ranges from giving basic knowledge to how to analyze various biological data according to commonly used analysis pipeline.

In the e-learning course, about 60 lectures on various topics from basics to advanced applications are available (as of January 2014), and new lecture videos are added every year. We hope you will join us.

For more details, please see our website at: http://hpci.cbrc.jp/

SCLS Gotcha!

# Short Lecture at the K computer Booth in SC13

The Supercomputing Conference 2013 (SC13) was held in Denver, United States, in November 18 to 22. The RIKEN HPCI Program for Computational Life Sciences sent two people to the conference: Atsushi Miyauchi, a research scientist of the High Performance Computing Development Team, and Erika Jinnai, a member of the Planning and Coordination Team, to collect information on supercomputers in the world.

SC13 is the world's biggest international conference on high-performance computing technology, and the year marked the 25th year since it was first held. It has provided opportunities for information exchange, discussion and collaboration, which help promote utilization of high-performance computing (HPC), and which are useful for research and education aimed at scientific development and greater knowledge.


Staff members explaining research activities at our K computer booth

In this conference, in addition to sessions and thesis presentations, the supercomputer ranking "Top 500" and the Gordon Bell Prize are announced, and exhibition booths of universities, research institutions and companies are set up.

The RIKEN Advanced Institute for Computational Science, which operates and manages the K computer, set up an exhibition booth as it did last year to demonstrate research activities using the K computer, and held short lectures by lecturers from related organizations (including the Strategic Programs for Innovative Research).

SCLS short lecture


SCLS held a lecture titled "Collaboration of Education and Outreach Activities" with a desire to share our activities and knowledge, and provide an opportunity to globally develop collaborative relations in joint research, education and outreach activities.

After the lecture, we received questions about education in bioinformatics and other subjects, and were able to share information with U.S. government agencies which were concentrating on outreach activities. In addition, the lecture helped us to have interactions with organizations which showed interest in our activities, as well as Japanese education and outreach activities. Through the lecture, we became aware of possibilities to expand our activities.

# News & Events

## First call for research theme proposals for access to the SCLS supercomputer system during FY 2014

We will provide the computational resources of the SCLS supercomputer system, which is compatible with the K computer, to biologists and bioengineers at no charge. We will entertain research theme proposals as follows.

- Application period: February 3 (Mon) - March 3 (Mon)
- Notification to successful applicants: Late March
- Period of use: April 1 (Tue) - March 31 (Tue) 2015

For more details, please visit our website (http://www.kobe.riken.jp/stpr1-life/).

## On-site classes in high schools

To promote education in the computational life sciences, SCLS offers classes on computational life science for high school students and teachers throughout Japan to foster understanding of the latest science and technology. These classes can be given in schools and social educational institutions. In addition, seminars can be presented in science and technology promotion events.

For more details, please visit our website (http://www.kobe.riken.jp/stpr1-life/), or contact the Planning and Coordination Group of the HPCI Program for Computational Life Sciences by e-mail (senryaku1@riken.jp).

Exercise using base sequence puzzle at Okayama Prefectural Tamashima High School

Hokkaido University -RIKEN Joint Symposium
Student 5-minute Session

## Computer Simulation of Life Science toward Future Medicine: Challenge of the High Performance Computers to Medical Applications

The symposium entitled "Computer Simulation of Life Science Toward Future Medicine: Challenge of the High Performance Computers to Medical Applications" was held jointly by the Creative Research Institution of Hokkaido University and SCLS at the Hokkaido University Conference Hall on Thursday, August 1, 2013. Researchers engaged in cutting-edge research on utilization of the K computer and invited researchers in biomechanics, biological systems, and physiology presented potential applications of high performance computing in medical technology and diagnosis.

We also sponsored a student workshop open to undergraduate and graduate students from across Japan, entitled "What kind of medical applications will you do research on if you could use the K computer?" We would like to express our gratitude to the six participating graduate students for delivering excellent short presentations.

Speakers:
(From left) Mr. Kenta Yahata (Saitama Univ. Graduate School), Mr. Sunao Tomita (Hokkaido Univ. Graduate School), Ms. Stephanie Nix (Tohoku Univ. Graduate School), Mr. Naoki Takeishi (Tohoku Univ. Graduate School), Mr. Tomohiro Otani (Osaka Univ. Graduate School), and Mr. Masataka Arai (Kyushu Univ. Graduate School)

Strategic Programs for Innovative Research:
Field 1 & Field 2 Symposium in Nagoya University

## Computational studies of complex biomolecular systems to understand the role of intermolecular interactions

Strategic Programs for Innovative Research Field 1: SCLS and Field 2: CMSI held an interdisciplinary symposium, focused on the potential of high performance computing for studies of biomolecular complexes to understand intermolecular interactions, at the Lecture Hall of Building IB, Nagoya University on Tuesday, December 17, 2013.

Although the symposium was targeted primarily at researchers studying biomolecules, presentations were made by scientists from diverse backgrounds. Active question-and-answer sessions provided ample opportunities for interdisciplinary discussion.


## Third Japan High School Championship at Kagaku no Koshien: Science Olympiad (March 21 (Fri)-24 (Mon))

SCLS will organize an exhibition booth at the Kagaku no Koshien (March 23 (Sun)): Science Olympiad, which is a contest that tests scientific knowledge and skills of student teams from high schools all over Japan.

Strategic Programs for Innovative Research Field 1

## Supercomputational Life Science

SPIRE (Strategic Programs for Innovative Research) is a MEXT program aiming to produce ground-breaking results in computer science technology by maximizing the benefits of the HPCI (High Performance Computing Infrastructure) centered on the supercomputer K computer, and encouraging developments in five research fields that need to be strategically addressed.
"Supercomputational Life Science" has conducted research with the mission of understanding and predicting life phenomena based on large-scale simulation and advanced data analysis, and to apply these results to design and implement medicine and medical care with its research.


# BioSupercomputing Newsletter Vol.10 2014.3



RIKEN HPCI Program for Computational Life Sciences

7-1-26 Minatojima-minamimachi, Chuo-ku, Kobe Hyogo 650-0047, Japan
Tel:+81-78-940-5692 Fax:+81-78-304-8785 http://www.kobe.riken.jp/stpr1-life



Issued: March 2014

RIKEN 2014-023