Integrating Cutting-Edge Genetics with Genetic Resource Information for Research and Society

KAWAMOTO, Shoko Associate Professor

Genetic Informatics Laboratory,
Department of Informatics.

Ph.D. (Science) (Osaka University). In addition to running the "National BioResource Project," which integrates the collation and supply of Japanese bioresources, Kawamoto conducts research on how to make the database of genetic resource information useful for society, mainly in the field of genetic medicine. She specializes in genome biology, human genetics, and life, health, and medical informatics. Concurrently, she is also an Associate Professor at the Database Center for Life Science, ROIS.

Kawamoto carries out a wide range of research, particularly on how to extract data from model organisms and apply such information to human genetic medicine. While specializing in mammals, Kawamoto leads a project to collect, preserve, and provide 31 bioresources to other researchers around the world. Furthermore, she acts as a Japanese node of the international network Global Biodiversity Information Facility, a network which collects "biodiversity information" in nature. Moreover, the challenge of applying fundamental research on model organisms to the rare disease research has also begun under the AMED program named IRUD-Beyond starting in FY2017. In the age of the genome, Kawamoto addresses how to accelerate progress in biology, how to contribute to society through biological research, and how to pioneer a new integrated field of biology. She drives each project with such passionate interest.

Development of databases *is* essential

Kawamoto started her research career as a molecular biologist when she was a university student. She looks back: "Because it was a time when genome research had just started, I had to have a point of view on biology through looking at thousands or tens of thousands of genes (giggling)." As is widely known, a tremendous amount of human genome information continues to be produced, in addition to the large amounts of genomic information available on model organisms, including *Drosophila*, which can be browsed by everyone. "In order to elucidate the function of a gene, it is extremely important to identify which elements are the same and which are different between different species by comparing the same gene between different organisms," says Kawamoto. In other words, we need to ask the appropriate questions of the enormous genomic datasets which have already been obtained to extract biologically meaningful information.

In 2017, Kawamoto took charge of a project called the "National Bioresource Project (NBRP)," which collects, stores, and provides to researchers bioresources of various model organisms, such as mouse and rice. She says, however, that she considers that there is an issue with the database because of the lack of horizontal connections within the database; in other words, genes from different species cannot be fully compared in the database. "For instance, when you're looking at a plant genome from the viewpoint of breeding, it is important to be able to distinguish between a gene which is evolutionarily common to most plant species and which gene is, say, unique to monocotyledonous plants such as rice, wheat, and barley? Without such annotations, even if breeding using this gene succeeds with one plant species, it is impossible to predict whether the success can be achieved for other species."

Another reason for this loss of connectivity also resides in the biology research papers themselves. Kawamoto collects scientific papers

containing research data using the resources of the NBRP. “A huge amount of scientific knowledge has been accumulated in papers. It is necessary to collect evidence in a comparable form in order to analyze the data computationally,” she says. However, papers often only describe the point that the researcher thinks is important qualitatively. There is no standard format of describing phenotypes in a manner that is comprehensive to some extent and shared by researchers. “Genetics is composed of the genotype and the phenotype. However, the phenotype could not be easily compared between two organisms before.”

Connecting rare human genetic diseases and *Drosophila*

Surprisingly, Kawamoto once kept her distance from bioinformatics research that she says “it has been my long-term project which I feel is so important.” This reluctance was because of her concern that “no matter how informative the database is, how can we use it for research and to help society?” While conducting research on genetic medicine as a genetic counselor, she came across the idea of “using model organisms such as *Drosophila* to help elucidate rare genetic diseases in humans,” and that made her think that “I could somehow contribute because I had an interest in a theme concerning disease models in mammals.”

When investigating a rare genetic disease using a next-generation sequencer, she says that there are situations in which the causative gene change can be identified for genetic diseases, of which only a few cases exist in the world. It is not difficult to obtain the treatment information of other patients if the disease affects tens of thousands of people. However, it is quite difficult to find two people who have the same change in a specific gene in the case of a rare disease. “Mouse has been a key animal for disease models. However, a mouse is a model organism that, relatively speaking, is expensive to study. Therefore, not many experiments can be done. I wondered if we could accelerate the experimental elucidation of rare genetic diseases in humans using a model organism with a short generation time such as zebrafish or *Drosophila*.” According to Kawamoto, such organisms are useful for research on rare genetic diseases and are expected to use understanding the etiology of the disease. “I foresee a scene where a physician explains evidence from *Drosophila* to the patient while showing them a paper on a medical website, and that sort of scenario will become increasingly realistic from now on.”

Connecting researchers of genetic medicine and model organisms by cooperating with the Japanese Agency for Medical Research and Development

In November 2017, a research project entitled “Japanese Rare Disease Models and Mechanisms Network (J-RDMM)” started with NIG as the representative institution under the Japan Agency for Medical Research and Development (AMED). Simply stated, it is “a project to investigate the cause of genetic diseases using model organisms.” NIG has long contributed to this field in terms of genetic resources and database development of model organisms. Kawamoto is one of the key working members of the project, in which NIG is demonstrating its research strengths.

Kawamoto describes her research strategy as follows. “I will call on model organism researchers nationwide and ask them to register to participate in research on intractable and rare genetic diseases to collaborate with clinical researchers. Because clinicians specializing in hereditary diseases are well aware of “which tissue should be checked at what time” to characterize the symptoms of the patient, we will ask them to submit a request for the model organism that they want to be analyzed based on the type of genes presumed to be responsible for the hereditary disease in question. On the other hand, I will request the model organism researchers to propose an efficient analysis method for the gene responsible, depending on the biological process they are studying. We began to create such collaborative groups, comprising medical researchers and model organism researchers in April 2018 and now, we have approximately 50 active collaborative groups. Our goal is to output some results for approximately 50 diseases in the coming year.”

Moreover, Kawamoto predicts a lot of applications for this strategy beyond genetic medicine. “In the near future, you may receive thousands of pieces of information from the results of your health examination. I hope we will develop fundamental technology that will allow us to examine and understand all these components of the life system in an integrated way. As a first step in the development of such a system, my interest in the bioresource work, that can connect diseases to organisms in various ways, never runs out.”

Genetics *is* important

While working on such a complex project, Kawamoto strongly feels that in the environment of NIG “Again, genetics is important.” “Genetic information no longer refers simply to individual genes but rather to integration of the information from the metabolomics and all the other -omic technologies derived from the genome. Without being able to analyze and understand the interactions associated with the genome in the form, say, of the genetic background and environmental factors, it is difficult to predict what effect a gene (and its mutants) would have even from an experiment performed in a laboratory under controlled conditions.”

“Bioinformatics” is a field in which life science and information science merge and in which Kawamoto has been involved since her university days, and the field has advanced dramatically in recent years. The whole genome of numerous organisms has been sequenced and the genes annotated and now, anyone can access and browse such data. As Kawamoto says: “That is why bioinformatics has to connect more research institutions and researchers and contribute to genetics through a heuristic use of knowledge from experiments using computation.”

Interviewer:
Rue Ikeya, Research Organization of Information and Systems

Photographer:
Mitsuhiko Kurusu, Office for Research Development

September 2018