BioResource Now !

Issue Number 8 January 2012

Kazuo Moriwaki (RIKEN BioResource Center)
At the End of the 2nd Stage NBRP

P1 - 2

# At the End of the 2nd Stage NBRP

Kazuo Moriwaki
Senior Advisor, RIKEN BioResource Center

Development History of BioResource Projects

I have been involved in the National BioResource Project (NBRP) and RIKEN BioResource Center (BRC) for the last 10 years. Before that, I participated in the strain preservation project of the National Institute of Genetics (NIG) for about 30 years. Dr. Kihara, the 2nd director of NIG, used wild wheat as the research organism. Since then, mainstream research at the NIG has focused on elucidating the principles of various biological functions in living organisms, as demonstrated by the research of Dr. Hikoichi Oka and Dr. Hiroko Morishima, who have been studying wild rice, and the research of Dr. Toshihide Yoshida, who has been studying the wild black rat. These research projects led to the collection, development, and preservation of strains of the specific organisms targeted by each project.

Research conducted by scientists is the driving force for the collection and preservation of resources, and this principle of strain preservation has been unchanged through today. The Jackson Laboratory, which has developed into a prominent mouse resource center, was also founded by researchers who began to collect mouse strains for their own research. As this example demonstrates, along with the development of each research field, demands from the research community for the use of resources have overwhelmed the individual efforts of researchers who developed these resources.
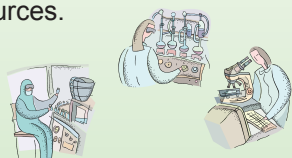
To fulfill the expanding need for mouse resources, the Jackson Laboratory has grown in the US, while many different resource centers in the EU nations have been developed. In 1951, more than half a century ago, the strain preservation fund was provided by the former Ministry of Education in Japan in order to meet the demand from pioneering researchers at the time; however, the budget for this fund was very small, even though the continuity of the fund was secured. Nevertheless, ample support for the mouse resource project was provided continuously for a long time due to the special research focus in cancer. In the 2000s, RIKEN BRC and NBRP were eventually founded, and the scale of bioresource projects in Japan became comparable to international resource centers.

Pioneer Spirit and Historical View of Bioresources

From the viewpoint that scientists' pioneering research is the driving force for consolidating bioresources, an accurate prediction of the direction of life science research is important for advancing bioresource projects. Because of the recent advancements in genome analysis technology, gene expression mechanisms can now be elucidated at the molecular level. In addition to the elucidation of genetic mechanisms, the scope of the research is now being expanded to include genes that regulate the expression of other genes and even the activation systems located upstream of these genes. The involvement of gene regulatory networks (GRNs), which are responsible for the expression traits of each gene, is also being elucidated.

Each gene comprising the network exhibits a long evolutionary history of mutations and adaptations to their environment, and each resource targeted by NBRP today can be considered a witness of this history.

Research in gene network mechanisms will advance markedly with the development of genome-level studies, and the resources used in this research, including the history of the genome, will serve as additional research materials. Some genetic properties of model resources have been affected by genetic selection, which was imposed in the process of resource development for the purposes of microbiological and genetic purifications, and thus, the quality of the resources as a model may not necessarily reflect the original "organism." For example, experimental mice, which have been widely used as model organisms, may potentially lack genes responsible for anticarcinogenesis and selective mating, which originally existed in wild mice.

Continuity of the Projects and Human Resources

Continuity is one of the fundamental principles of administrating bio-resources, and because of this, human resources are the most important component. In particular, it is crucial to hire researchers who have an excellent record of research achievements using bioresources and thoroughly understand the significance of bioresource projects. It is desirable that prominent researchers take the lead in administrating resource projects and reliable technicians undertake practical tasks.

A number of resource projects that have been administrated actively for a long time were discontinued immediately after the disengagement of primary researchers, and precious genetic resources were lost. In order to secure the continuity of resource projects, it is crucial to recruit researchers to lead the projects and to train reliable technicians.

While there are defined career paths for researchers, technicians do not generally have planned career paths. It is not easy for technicians to be promoted by publishing academic journal articles, and this is not necessarily ideal either. Hence, independent career paths should also be developed for technicians. Nevertheless, the understanding and cooperation of

researchers are keys for developing technicians in the life sciences and securing a stable research environment. In general, one can gain the understanding of researchers easily enough, but both players may in fact need to share the same pie. Bioresource projects thus need, above all, the support and cooperation of researchers' communities.

Easy Data Acquisition by Ebot

Ongoing Column [No.65] 10min

Here, I will introduce a method through which a keyword search in PubMed can be conducted and the search results can be obtained in XML using the Ebot service provided by the National Center for Biotechnology Information (NCBI). Ebot is an interactive service that automatically generates and provides Perl script according to the input content. The Perl execution environment needs to be installed in Windows, but not in Mac or Linux operating systems, where scripts can be used immediately. Let us begin using Ebot!

Installation of the Perl execution environment (only in Windows)

Download and install the Perl execution environment from ActivePerl. Select "Windows x86" if a 32-bit version of Windows is used, and select "Windows 64-bit, x64" if a 64-bit version is used.

Active Perl: http://www.activestate.com/activeperl/downloads

Fig. 1: Download window for ActivePerl

Ebot

Input data by following the input window in Ebot. In this example, we will write a script to obtain data from PubMed that match arbitrary keywords. Input codes for each window (※) are summarized at the end for reference.

Ebot：http://www.ncbi.nlm.nih.gov/Class/PowerTools/eutils/ebot/ebot.cgi

First, an input method for the search query and a search target server should be selected (※1). Regarding the search query method, users can select either to directly specify the search query or to upload a file for query (Fig.2- ①). Only one search target server can be selected per search. Users can write scripts that target not only PubMed, but also the other databases such as "Protein," "Nucleotide," and "EST" (Fig.2- ②).
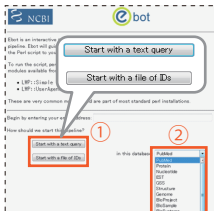
Fig. 2: Top page of Ebot

The search query and search limits are then specified (※2). For the search query, keywords should be input (Fig.3- ①) and the period can be specified if necessary (Fig.3- ②). If not specified, data for the entire period will be targeted.

Fig. 3: The search query and search limits settings page

The output content for the search results is then selected (Fig. 4 A)(※3). Summaries, full records, etc. can be selected. Subsequently, the output file and format in which search results are saved are specified (Fig. 4 B) (※4). Select HTML if the downloaded data will be viewed by a browser, and select XML if the data will be edited.
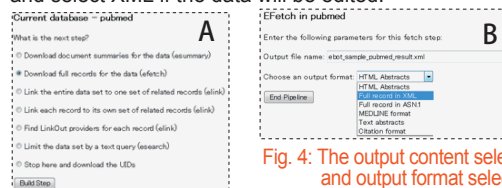
Fig. 4: The output content selection page (A) and output format selection page (B)

Finally, the script name is specified (※5). The script will be output in Perl, and thus, add the ".pl" extension to the script name (Fig. 5-①). Click "Generate Perl" (Fig. 5-②), and complete the task by downloading the file.

Generate Perl

Fig. 5: Script name setting and file generation

How did you like script writing in Ebot ? If the execution environment of Perl has been installed, the script can be run immediately to obtain data (※6). In addition, the generated script can also be expanded by your own so that the script will meet various needs.

If you feel that it is tiresome to obtain data one by one by opening the NCBI website in a browser, please utilize this service.

（Tohru Watanabe）

Input content for each page

※1: Selection of an input method for the search query and a search target server
　Search target server: "In this database: PubMed"
　Input method for the search query: "Start with a text query" button
※2: Selection of search parameters
　Search query: "Text query in PubMed: NBRP"
　Search limit: "Retrieve records for all dates"
※3: Settings for output content
　Output content: "Download full records for the data (efetch)"
※4: Settings for the output file and output format
　Output file: "Output file name: ebot_sample_result.xml"
　Output format: "Choose an output format: Full record in XML"
※5: Settings for the script file
　Script name: "Enter the filename for your Perl script: ebot_sample.pl"
※6: Execution of the script file
　Open a command prompt and input "perl ebot_sample.pl."
　Then, "ebot_sample_result.xml" will be output in the same folder as the script file.

Recommended Book ! 〈NO.1〉

"Epigenetics: The Ultimate Mystery of Inheritance"
by Richard C. Francis (translated into Japanese by Kyoko Nonaka; Diamond, Inc.)

This book is probably the first independent book on the topic of epigenetics for public readers.
Epigenetic alteration is the regulation of gene expression caused by various modifications to chromosomes (DNA and protein), which play a role as "information media," without altering the genetic information (nucleotide sequences) encoded in the DNA. Therefore, these epigenetic alterations will not necessarily be inherited by descendants (although there are some exceptions). Epigenetic alterations occur along with cell differentiation, and these alterations will be copied and inherited by daughter cells after cell division. As a result, these alterations are directly related to ontogenesis and the onset of diseases such as cancer. Each case example is narrated episodically with quotes from original articles.
According to the author, epigenetics urges us to rethink conventional genetic concepts, such as the "gene versatility theory." So far, there are several interesting books on the subject of "genes or environments," such as The Century of the Gene, by Evelyn Fox Keller and Nature via Nurture; Genes, Experience and What Makes Us Human, by Matt Ridley. However, Epigenetics: The Ultimate Mystery of Inheritance provides such a significant impact that these other books on genes and genetics may seem utterly old-fashioned. This suggests that research in the life sciences is advancing rapidly. (K.N.)

Editor's Note

Dr. Moriwaki, who is considered the father of NBRP, contributed an article in this January issue, in the year when the 3rd Stage of NBRP will begin. His enthusiastic and recondite thoughts on bioresources are condensed in his short sentences, yet his words do not sound intimidating to readers, which may be attributed to his character. We sincerely appreciate his contribution. The "Recommended Book" is a gift from one of my senior associates. This year is off to a promising start. (Y.Y.)

BioResource Information

(NBRP) www.nbrp.jp/
(SHIGEN) www.shigen.nig.ac.jp/
(WGR) www.shigen.nig.ac.jp/wgr/
(JGR) www.shigen.nig.ac.jp/wgr/jgr/jgrUrlList.jsp

BioResource Now !
■ ■ Issue Number 8 January 2012

"translated by ASL translation service"