Issue Number 3 August 2007

BioResource now !

Our monthly newsletter features a variety of information, highlighting current domestic and international issues concerning bioresources.

Introduction to Resource Center No.19

National BioResource Project "Medaka"

Kiyoshi NARUSE
Laboratory of BioResource, National Institute for Basic Biology

Download the PDF version of this newsletter at

http://www.shigen.nig.ac.jp/shigen/news/

Other information on bioresources is available at

NBRP	http://www.nbrp.jp/
SHIGEN	http://www.shigen.nig.ac.jp/indexja.htm
WGR	http://www.shigen.nig.ac.jp/wgr/
JGR	http://www.shigen.nig.ac.jp/wgr/jgr/jgrUrlList.jsp

"2nd International Biocuration Meeting"
October 25–28, 2007 at Dolce Hayes Mansion, San Jose, CA, USA
Details are available at the following website: http://biocurator.org/Mtg2007/

Cell Technology initiated a new series entitled " Let's Use! BioResource ". Bioresources in Japan will be introduced in a series of 14 volumes from November 2006 to December 2007.

Introduction to Resource Center No.19

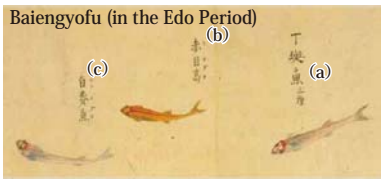
National BioResource Project "Medaka"

http://www.shigen.nig.ac.jp/medaka/

Kiyoshi NARUSE, Associate Professor
Laboratory of BioResource, National Institute for Basic Biology

In Japan, medaka (killifish, *Oryzias latipes*) is one of the most familiar fishes. Himedaka (the orange-red variety) and Shiromedaka (the white variety) have been reported as early as the Edo Period and have been illustrated on Ukiyoe (Japanese wood block prints) as a characteristic of summer (Fig. 1). For the Japanese, medaka is a cherished and favorite pet that well represents an idyllic environment and it is used by researchers in Japan as an excellent experimental animal. Mendelian inheritance of medaka was confirmed in 1916, and its sex-limited inheritance was reported for the first time by Dr. Aida in 1921. Its artificial sex transformation reported by Dr. Yamamoto in 1953 was also the first successful example in animals. As a result, medaka has been used in extensive fields such as genetics, physiology, endocrinology, embryology, and environmental science. Due to such research activities, a large genetic diversity has been observed in the natural groups of medaka, and the supposed inbred strains, which are genetically homogeneous, have been established. Research for exploring the ancestors of medaka has contributed to the biogeography of Asia. The development of an efficient gene transfection technology using medaka is also an important advancement.

"Medaka Scooping" by Harunobu Suzuki,



From the left, Shiromedaka (the white variety), Himedaka (the orange-red variety), and the wild type
(a) Medaka
(b) Akamedaka (the orange-red variety)
(c) Shiromedaka (white variety)

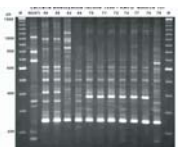
The orange-red medaka variety is known as Himedaka.The white variety is called Shiromedaka. Kowagyofu (in the Edo period)

Fig. 1: Medakas illustrated in the Edo period Himedaka and Shiromedaka were described as early as the Edo period (courtesy, Hori's laboratory, Nagoya University).

In this issue, we will introduce the construction of a linkage map that covers the whole genome, a large-scale EST analysis, the development of mutant strains using N-ethyl-N-nitrosourea mutagenesis (ENU), and the medaka genome analysis project, all of which were initiated around the year 2000. Additionally, we will discuss the significance of the Medaka BioResource Project as the central repository of medaka resources, which have been developed through these researches.

1 Construction of a Genome-Wide Medaka Linkage Map

For a species that can be genetically analyzed, the construction of a linkage map that covers the whole genome by identifying as many numbers of linkage groups as that of chromosomes is considered an extremely important step to establish the species as an experimental animal. Such a linkage map of medaka was constructed by a multipoint linkage analysis using DNA polymorphism.



Genetically homogeneous inbred lines of medaka have been developed, and genetically and significantly distinct regional groups have been identified. Therefore, the efficient construction of a linkage map was considered possible once an effective method of identifying DNA polymorphisms was established. The RAPD method was the first to be used for identifying DNA polymorphisms. The first linkage map constructed using this method was reported by Wada, *et al.* in 1995. Subsequently, the first genome-wide linkage map containing 24 linkage groups was constructed by Naruse, *et al.* in 2000 by using the AFLP method and the PCR-RFLP method, which identifies polymorphism by amplifying specific regions by PCR and treating them with restriction enzymes. Currently, an extremely accurate linkage map harboring approximately 2500 gene loci is being constructed using SNP polymorphisms.

2 Transcriptome Analysis of Medaka by EST Analysis

For establishing an organism as an experimental model, a catalog of the transcripts of that organism needs to be prepared. One of the methods for preparing such a catalog is EST analysis. With this method, one can prepare a catalog of transcripts that are expressed in each developmental stage or tissue by constructing a cDNA library and sequencing the 5′ and 3′ ends of randomly selected clones from the library.

In 1998, Naruse, *et al.* initiated the EST analysis using an inbred line of the northern Japanese population of medaka (HNI). Almost simultaneously, a large-scale EST analysis was conducted primarily by the National Institute of Genetics by using the Hd-rR/d-rR strains. Currently, approximately 220,000 sequences have been registered in DDBJ, EMBL, and GenBank, and there are approximately 39,000 unique sequences. The Medaka BioResource Project can distribute approximately 130,000 clones (more than 13,000 unique sequences).

Thirteen new research subjects for the 2nd stage of the NBRP have been announced.
With regard to biological species, *Ciona intestinalis*, tomato, *Bacillus subtilis*, cellular slime molds, and general microbes of RIKEN BRC are the new additions to the NBRP.
Please refer to the NBRP website for further details.
URL: http://www.nbrp.jp/

3 Construction of Artificial Mutant Strains by ENU on a Large Scale

Molecular mechanisms underlying numerous unknown biological phenomena have been uncovered by extensively constructing a mutant strain whose certain phenotypes follow Mendelian inheritance. Further, the abovementioned molecular mechanisms have also been revealed by identifying the genes responsible for these phenotypes. Regarding medaka, a genome-wide linkage map and a BAC library were constructed and projects that extensively collect mutant strains were conducted primarily in Japan to correlate with the advancement of EST analysis. Among these, Dr. Kondo's ERATO project which was initiated by the Kondo group in collaboration with internal and external researchers was the biggest project for artificial mutant strains development. This research project established more than 480 mutant strains. Besides this project, the identification of mutant strains with abnormalities in regeneration, somitogenesis, bilateral symmetry, and organogenesis of endodermal organs has been enthusiastically conducted. As a result, causative genes have been identified in series by using the ample genome resources (Fig. 2).

Fig. 2. Example Mutant Strains of Medaka
Numerous mutant strains exhibiting specific phenotypes have been isolated. *Fgfr1* was reported as a causative gene of *headfish* by using a positional cloning method (courtesy, Takeda's laboratory).

The history of mutant strain identification is known today only because of the long-term researches on natural mutant strains conducted by Dr. Tomita in Nagoya University. Dr. Tomita crossed strains that harbor rare mutant genes that are found in natural groups over 3 generations to construct homozygotes, and found mutant strains that exhibit extremely unique phenotypes. More than 70 strains have been identified. Positional cloning in the early stage has been conducted mainly by using these strains. Intriguingly, most of these natural mutation strains were developed using transposons. In this case, the importance of transposons as the cause of mutations is directly evident. *Tol2*, which is a DNA-type transposon of medaka, is an effective and important tool to construct transgenic zebrafish, a compact fish analogous to medaka.

4 Medaka Genome Analysis Project

A project that attempted to completely sequence the medaka genome by using the whole genome shotgun method was initiated in the fall of 2002; it was a challenging task at that time. This project was initiated as a collaborative project between the National Institute of Genetics and the University of Tokyo. Subsequently, RIKEN, National Institute of Informatics, Niigata University, and Keio University also participated in the project, and it thus became a nationally promoted project.

Obviously, the large-scale screening project of mutant strains described in the previous section had been planned at the time. Positional cloning is a method to identify the causative genes of mutant strains. In this method, a detailed linkage map containing gene loci of mutant strains is constructed by using DNA markers, and subject or candidate genome regions are physically reproduced with the BAC clones by screening the BAC library. Next, the causative and prospective genes of mutant strains are identified by sequencing all the base pairs of the BAC clones. Further, the base pairs that differ between mutant and wild-type strains are determined by comparing the prospective causative genes.

This method is extremely time-consuming and costly. If the sequence of a whole genome has already been determined, genes contained in the detailed linkage map of mutant strains can be immediately identified during map construction. After this stage, the only tasks that remain are sequencing of the genes of the mutant strains and comparison of the sequences; these enable the identification of the bases and the genes that are mutated. The research trend would transit from mutant strains to causative genes in an accelerated speed once the whole genome sequence is determined.

Fig. 3. Medaka Genome Browser (UTGB)
Information such as base sequences, transcription initiation sites, and links to BAC and Fosmid can be obtained.

Medaka Genome Analysis Project, which was initiated in the fall of 2002, released the first genome assembly data (http://medaka.utgenome.org/) in 2004 (Fig. 3). A draft-level analysis was completed in 2006, and the detailed analysis result of medaka genome with focus on genome evolution was successfully published in Nature in June 2007. The success of this Medaka Genome Analysis Project has indicated that sequencing of the genome of a vertebrate, annotating genes, and developing a platform to release the data are possible even by the researchers of our country.

5 Aims of the 2nd Stage of the Medaka BioResource Project

The 2nd stage of the bioresource project was initiated in April 2007. Three institutes—National Institute of Basic Biology (core institute), Niigata University (sub-institute), and National Institute of Radiological Science (sub-institute)—were selected for the implementation of the 2nd stage of the bioresource project. During this stage, we have started building a comprehensive system to distribute the 464 strains of live resources (Fig. 4), clones derived from transcripts (cDNA and EST clones), genome resources including genome-fragment clones such as BAC and Fosmid constructed during genome analysis, and provide in silico information such as genome base sequences.

One of the aims of the 2nd stage of the Medaka Bioresource Project is to function as a central repository. In the 1st stage of the bioresource project, the institutes that developed and provided resources were almost the same; therefore, the function of the project as a repository was not particularly focused upon.

Fig. 4. Medaka Strains Distributed by the Medaka BioResource Project
Various types of strains such as general, mutant, transgenic, local wild type, and inbred strains can be provided.

Since the National Institute of Basic Biology was declared the core institute for the 2nd stage of the project, developers and providers of most of the medaka bioresources were separated. Although this seems to be disadvantageous, this separation would enable the provider (i.e., institutes implementing the bioresource project) to mediate various conflicts of interests that would possibly occur between developers and users. Moreover, if this system is well executed, it would lead to a cycle of resources, in which the existing bioresources would be used to develop new resources, and these new resources would be deposited to the institutes participating in the bioresource project and would be in turn provided to develop more new resources.

Undoubtedly, the speedy distribution of these well-managed bioresources is also important. The occurrence of various issues such as the conflict of intellectual rights is anticipated in the distribution of resources. We would like to hasten the distribution of bioresources as much as possible by making a preliminary clarification, resolving these issues, and digitizing the conclusion of MTA (a contract defining the conditions of using strains). We profoundly hope that more and more researchers use Medaka as a research resource, and we believe that their findings will lead to novel understandings of biological phenomena.

The 2nd stage of the project is at a primitive stage, and a sufficiently effective system of collecting, preserving, and providing medaka bioresources has not yet been prepared. Nevertheless, we would like to gradually but steadily consolidate the system. Would you not like to use Medaka?

Editor's Note: After receiving the baton from Dr. Wakamatsu of Nagoya University who conducted the 1st stage of the NBRP Medaka, Dr. Naruse commenced activities of the 2nd stage of the NBRP in April 2007. Dr. Naruse played a central role in the Medaka Genome Analysis Project, and we feel that his strong enthusiasm and passion cannot be justifiably described in words. The establishment of inbred lines and the continuous preservation of resources have greatly contributed to the success of Medaka Genome Analysis. (Y.Y.)

Contact Address:
1111 Yata, Mishima-shi, Shizuoka 411-8540, Japan
Center for Genetic Resource Information, National Institute of Genetics
Tel: 055-981-6885 (Yamazaki)
E-mail: BRnews@chanko.lab.nig.ac.jp

"translated by ASL translatiaon service and proofread by Sharoh Yip"