BioResource now !

July 2005

BioResource now! No.7 is here

Download the PDF version of this newsletter at
http://www.shigen.nig.ac.jp/shigen/news/news.jsp

Hot News from Abroad (Gene Ontology) No.6

Gene Ontology

Gene ontology attracts attention

In recent years, the term “gene ontology (GO)” is being used by biologists and bioinformaticians. The familiarity of this term differs considerably among people, i.e., some may be well versed with the term, while others might have recently heard about it, or are hearing it for the first time now. Although I majored in life science, I was not acquainted with this field until a year ago. People who have heard about GO for the first time can look up this term using search engines, such as Google. The popularity of GO can be estimated from the large number of hits obtained.

About GO

It is beyond imagination to ascertain the number of genes being studied by numerous researchers world wide; however, these numbers are expected to be very high. Since numerous papers are published to report the progress of research technology, identifying the paper of interest from a stack of papers is not easy. Further, the writing style of each paper may result in difficulties in understanding the content correctly. Typically, when a paper on a certain gene is published by a certain researcher, other researchers who study the same species can understand the content without much difficulty. However, researchers who work on other species or multiple species as well as database developers who compile data and provide them to the public find such studies difficult to understand. The writing style, which is typical of a peculiar field of study, becomes an obstacle when researchers who study other species try to obtain correct data. Even database developers struggle to collate gene data of different living organisms.

In order to resolve such issues, GO provides the concept of accessing gene information by using an ID number; the gene information is presented in a special vocabulary.

GO has a hierarchical structure that encompasses three GO levels, namely, the biological_process, molecular_function, and celluer_component as roots (refer to the following figure). The lower GO level inherits the properties of the higher GO level. This hierarchical structure is a characteristic of GO, and users (e.g., researchers and annotators) can assign the correct GO to the target gene by considering the low and high hierarchical relationships of the gene (Note).

Moreover, a person who has not classified the gene can determine its characteristics from the hierarchical structure. Furthermore, the vocabulary provided by GO is basically defined by the gene, and it is not restricted to a frame of species, althou gh some exceptions do exist. GO can play an important role in the integration of gene data of various species.

(Note) Although GO provides the vocabulary together with the ID for the gene of interest, the vocabulary may not necessarily be used in research paper. A detailed expression is often required for explaining the characteristics of the gene more precisely. The important point is that the gene is described by the assigned GO; this GO does not restrict the manner in which the gene data is expressed. Moreover, a one-to-one correspondence between the gene and GO is not necessarily.

The hierarchical structure of GO can be represented as follows.
Example of higher hierarchy

```
GO : 0008150 biological_process
        GO : 0007610 : behavior
        GO : 0000004 : biological process unknown
        GO : 0009987 : cellular process
                GO : 0007154 : cell communication
                GO : 0030154 : cell differentiation
                  ・
                  ・
        GO : 0007275 : development
        GO : 0040007 : growth
          ・
          ・
GO : 0005575 molecular_fanction
        GO : 0005623 : cell
        GO : 0008372 : cellular component unknown
        GO : 0031012 : extracellular matrix
          ・
          ・
GO : 0003674 celluer_component
        GO : 0016209 : antioxidant activity
        GO : 0005488 : binding
        GO : 0003824 : catalytic activity
          ・
          ・
```

Gene Ontology Annotaiton Camp

The Gene Ontology Annotation Camp was held for four days from June 1 to June 4, 2005 at the Clark Center, Stanford University, California, USA. Although the lot size of universities and research organizations are comparatively large in Japan, I was overwhelmed by the area of the campus of Stanford University. Approximately 50 people from USA, Canada, Britain, Japan and other countries participated in the camp. These included not only biologists and bioinformatists but also those who studied information archives.

The Camp began with an introduction to GO, and explanation of the usage of GO in well known databases such as MGI, UniProt, DictyBase, ZFIN, TAIR, TIGR, and WormBase. This was followed by paper presentation and exercises that involved assigning GO. For the exercises, all participants were divided into groups of approximately five people. Two groups were made to read the same paper and assign GO; they then crosschecked their results with each other. Ideally, both groups should have the same result but different results were reported as it was a manual task after all. However, the hierarchical structure of GO was found useful here. Even when two GOs differ in their hierarchy, the distance of their separation was limited. Since the lower GO level inherits the properties of the higher GO level, the results can be summarized based on the higher GO level. In some cases, a lower GO level is more appropriate after the entire paper is read. If the GO level cannot be assigned with certainty, GO of a higher hierarchy level should be assigned. In any case, the entire paper should be read carefully first.

（Stanford University）

The campus of the Stanford University extends across 8200 acres (one acre is approximately 4047 m 2).
I traveled around the campus using the complimentary bus service.

（Palm Drive）

As the name suggests, palm trees line both sides of the street in the university.

Fig.1 Part of the GO annotation exercise performed during the Camp

The following is an abstract of a paper entitled “Functional Analysis of the Subunits of the Chromatin Assembly Factor RSF.” which was published in Molecular and Cellular Biology, October 2003, pg. 6759-6768, Vol. 23, No.19

6762: Rsf-1 is a nuclear protein

Immunofluorescence experiments in HeLa cells with highly specific antibodies to the RSF subunits Rsf-1 and hSNF2H (Fig. 2A) showed that the large subunit of RSF and hSNF2H localized to the nuclei (Fig. 2B),

Rsf-1 GO:0005634 IDA

hSNF2H GO:0005634 IDA

correlating with its purification procedure, which began with a HeLa cell-derived nuclear fraction. The staining pattern obtained with the antibodies appears to be specific; a similar pattern was obtained with antibodies directed against different regions of Rsf-1. Interestingly, during mitosis, Rsf-1 exhibited a more diffuse localization, whereas hSNF2H remained associated with mitotic chromosomes.

hSNF2H GO:0000793 IDA

The experimental results of two genes, Rsf-1 and hSNF2H , published in the paper are shown in Fig. 1; Their assignments are as follows: Rsf-1 GO: 0005634 IDA, hSNF2H GO: 0005634 IDA, and GO: 0000793 IDA (this example is based on the abstract of the paper; however, the entire paper needs to be read when assigning the GO). Moreover, the term IDA that follows the GO number should be noted; this is the Evidence Code, which is an index of the reliability of the GO. IDA is the abbreviation for “inferred from direct assay” and it implies that the result was obtained by direct assay. There are a total of 13 Evidence Codes together with IDA.

Evidence Code
IC: Inferred by Curator
IDA: Inferred from Direct Assay
IEA: Inferred from Electronic Annotation
IEP: Inferred from Expression Pattern
IGI: Inferred from Genetic Interaction
IMP: Inferred from Mutant Phenotype
IPI: Inferred from Physical Interaction
ISS: Inferred from Sequence or Structural Similarity
NAS: Non-traceable Author Statement
ND: No biological Data available
RCA: inferred from Reviewed Computational Analysis
TAS: Traceable Author Statement
NR: Not Recorded

Details on the Evidence Code can be accessed from

http://www.geneontology.org/GO.evidence.shtml#hier

Based on my experience at the meeting, I feel that GO will become popular henceforth. The sponsors and presenters of the meeting appeared to be enthusiastic about promoting GO. During the question and answer session, a very active discussion was held among participants, which is uncommon in Japan. As a database architect, I have high hopes for GO which enables integration on genetic level, irrespective of their species. Please use GO.

（Clark Center）

Clark Center with its modern architecture. This is the site where the Camp was held.

（Stanford city ）

Stanford city with beautiful houses

GOALL http://shigen.lab.nig.ac.jp/ontology/
Oryzabase http://shigen.lab.nig.ac.jp/rice/oryzabase/top/

Details on GO can be obtained from the Gene Ontology Consortium website: http://www.geneontology.org/

（Author: Shingo Sakaniwa, Genetic Information Laboratory, Center for Genetic Resource Information ）

Resource Feature (Guide Dogs) No. 5

A Gene Bank of Superior Guide Dogs

A need for more guide dogs

I would like to introduce a gene bank of guide dogs. Although the training scene of guide dogs, the active role played by these dogs, and the care of retired old dogs have been introduced on television shows, the use of a gene bank for guide dogs has never been featured. I have been thinking that the reason for the insufficient number of guide dogs is simply due to a shortage of funds.

At present, there are 900 guide dogs in Japan. However, the number of people requiring guide dogs exceeds this number by five times and there are 10 times more people who wish to own a guide dog. In addition, only 30 percent of the dogs which undergo strict training manage to pass the aptitude test necessary for them to become guide dogs.

Whether or not the inborn characteristics of a dog meet the conditions required for a guide dog is also an important factor. Since the dog's training process only begins after its contraception or castration, a guide dog cannot reproduce even if it is found to be of a superior breed. Therefore, a dog can only be either a guide dog or a breeding dog, disabling a guide dog to be used as a breeding dog at a later stage. If the number of breeding dogs decreases, the offspring will become genetically closer, and this may cause various problems in the future. At present, it is also difficult to import guide dogs from abroad.

Reproductive engineering promotes the breeding of guide dogs

There are research groups trying to overcome the shortage of guide dogs with reproductive engineering. Professor Hiroshi Suzuki of the National Research Center for Protozoan Disease, Obihiro University of Agriculture and Veterinary Medicine is a member of one of the leading groups. Together with Mr. Suwa, a team leader of the Hokkaido Guide Dog Association, he is promoting a project entitled "Reproduction engineering research on breeding of superior guide dogs" which is part of the "Promotion Subsidy for Science and Technology for Important Problem Solution Type Research".

This project aims to establish "K9 Biobank" (the name is derived from the English word canine)" wherein the genetic resources (sperms, embryos, ovaries, etc.) of superior guide dogs are frozen and stored (Photograph 1), and the reproduction of superior guide dogs can be realized, if needed. If the freezing-defreezing technology of sperms, embryos, and ovaries can be applied, long-term storage and transportation of genetic resources will become easy, and it would be possible to import them from abroad.

Professor Suzuki played an important role in the reproduction engineering of mice and is applying the technology developed with mice to dogs. However, he is facing difficulties because of the differences between dog and mouse strains. Reproduction engineering technology is very important for reproducing individuals from frozen resources. Although the artificial breeding of dog is regarded to be difficult, Professor Suzuki and his group succeeded in breeding a dog artificially by using frozen sperms two years ago. The artificial breeding using frozen ovaries will also be accomplished soon. Moreover, they will adopt a modern technique that uses genomic information to develop the parent-and-child discriminating method based on a satellite marker.

Photograph 1: A liquid nitrogen tank for storing frozen sperms

Construction of the database has started

We, the Center for Genetic Resource Information, National Institute of Genetics, are given the honor of participating in the K9 Biobank database construction which started this year. In the database, the basic information on embryos and sperms from superior guide dogs will be stored in the bank, and information on individual dogs (bloodline, case history, existence of etiologic gene, character evaluation result, and SNP) will be collected. We plan to construct a system that facilitates the sharing of information between related organizations and consequently fulfill the requests from the users.

The Asia GuideDogs Breeding Network (AGBN) was established in 2002 with eight facilities in Japan, one in South Korea, and two in Taiwan. The construction of this database is expected to lead to further cooperation among these facilities.

A visit to the Hokkaido Guide Dog Association facilities

I visited the Hokkaido Guide Dog Association in Sapporo on the day after the report meeting of the project at the Obihiro University of Agriculture and Veterinary Medicine was conducted. I toured the facility and was able to experience walking with a guide dog as a visually impaired person would.

During my visit, I came across some IT equipment that has been designed for visually impaired people; the equipment is used for enhancing web accessibility, which is an important field these days. A voice response web browser and color adjustment for weak-sighted people were among the IT equipment. According to the data obtained from the Institute for Information and Communication Policy, Ministry of Internal Affairs and Communications, 81.1% of the hearing-impaired and 69.7% of the visually impaired use the Internet. The usefulness of the Internet for handicapped persons can be understood based on the Internet utilization rate in Japan which was a high 60.6% in 2003. Therefore, equipment for enhancing web accessibility is of the utmost importance.

Photograph 2: PR dogs

The dog I walked with is a "PR dog" (Photograph 2). PR dogs are either dogs that used to work as guide dogs but do not do so anymore due to certain reasons or dogs that did not pass the strict evaluation test. After wearing the eye mask, I held the harness which is similar to that of a horse and fearfully entered an unknown world. I was pulled by the PR dog and began to walk somehow, although I felt that I must have walked quite strangely. The walk was quite fast. I concentrated all my energy on my ears and soles and walked in fright, feeling every bump in the road. Although the guide dog forewarned me about the obstacles, it was very difficult to know the types of obstacles. I truly felt how important the communication between a dog and a human was during this walk which lasted for about 30 minutes. Dogs have personalities and feelings, and there is also a matter of compatibility. If you want to be guided properly, it is necessary for you to stay with the dog for a long period in order to improve mutual understanding. (Photograph 3)

Photograph 3: Active guide dogs in a temporal stay cage

Photograph 4: A retired dog being cared for

It was very impressive to see a retired old dog being cared properly (Photograph 4) .I was also surprised to learn that local volunteers carry out many tasks such as rearing 2-month old dogs for 10 months, rearing the breeding dogs, and caring for retired dogs.

In future

Although the tenure of this project is three years, I would like it to be extended until it develops into a system that fulfills the needs of its users. Although the quality of Japanese guide dogs is very high, due to lack of a systematic approach to activities such as supplying guide dogs to visually impaired persons, frozen sperms circulation, and the use of fertilization technology, Japan is still behind the western countries in this area. The results of this project are eagerly awaited because many features required in a guide dog are common to the entire class of assistance dogs such as service dogs and hearing dogs. (Author: Yukiko Yamazaki)

Take a break...

- Do You Know About Google Earth? -

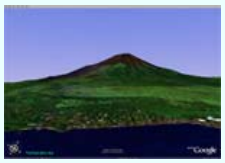
A virtual globe software named “Google Earth” (the trial version: http://earth.google.com/) has been launched by the renowned search engine Google. I tested it because it has a reputation of providing beautiful satellite photographs and having interesting functions.

The free version of Google Earth can be downloaded from the abovementioned URL and installed. When the program starts, a picture of the earth floating in the universe appears.

Let us check out New York. The software is simple to handle. The functions are mouse enabled. Left click controls the up-down and left-right movements, while right click is for zooming in on the image. Clicking the scroll button while moving the mouse around changes the angle at which the image is being viewed.

The Statue of Liberty can be seen in the satellite image. The photograph is so clear that even the number of trees in the park can be counted. If you look carefully, the shadow of each person in the photograph can even be seen.

A 3D display of buildings is also possible. This is a side view of Manhattan Island. The shape and size of each building is finely reproduced.



A 3D display of the height difference of geographical features is also possible through these images. This is an image of Mt. Fuji viewed from Kawaguchi Lake. Since this is the English edition, please do not pay attention if the display is marked as “Huji San” or “Fusiyama.”

After using this software, I feel invigorated; it appears as though I have flown around the world and visited many famous sightseeing spots within a short period of time. The commercial version of this software can be used for various purposes because it can be integrated with other systems such as those used in car navigation. (N. K.)

Information Technology Vol. 7

“Types of Search Engines - 2”

Although I casually mentioned search engines (SEO) in general earlier, they can actually be classified into two types of search engines.

- Robot type search engine

  The robot type search engine is a search engine in which a “robot” patrols the Internet and collects information for automatic searching.

- Directory type search engine

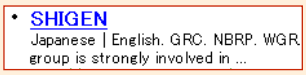
  The directory type search engine is a search engine in which homepages are manually sorted and classified into categories.

The advantages and disadvantages of both search engines are summarized in the table below.

Type	Advantage	Disadvantage	Typical SEO
Robot type	Large number of home pages can be searched	Since the number of hits is high, some degree of familiarity is required to search for the target page	Google, infoseek
Directory type	Accuracy of search results in this case is higher than that in the robot type because of manual judgment	Unregistered pages cannot be found	

As shown in the table above, each search engine has its own set of advantage and disadvantage. Therefore, a “hybrid type” search engine, which combines the advantages of the directory type and robot type search engines, has been gaining popularity recently.

The Yahoo! search engine first looks through pages with registered search keywords and displays the results if any exist. If there are no results, it will to look through the results obtained from the robot type search engine.



Robot type

Directory type

I will introduce methods to improve the search engine ranking of your homepage in the next issue. (Reference: Perfect plan to increase homepage access)

(Author: Takehiro Yamakawa of Genetic Information Laboratory, Center for Genetic Resource Information)

Editor’s notes: I was rather tensed on the first day of the Camp because it was the first time that I was participating in a meeting held abroad. While listening to the intense discussion among the participants, I felt that my passion towards my job was inadequate. There were three other Japanese participants, and after the Camp, they kindly guided me around San Francisco. I am extremely grateful to them. (S. S.)

It has been ten years since I began working on genetic resources. I entered an unknown world, the mysteries of which are still being unraveled. After participating in this project, I became aware of the present situation of the gene bank for guide dogs. It is a subject of high priority. (Y. Y.)

Contact: Yata 1111, Mishima, Shizuoka 411-8540, JAPAN
Center for Genetic Resources 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”