New trend of Biodiversity Informatics based on Global Biodiversity Information Facility (GBIF) and Barcode of Life (BoL)

Yasumasa Shigemoto1), Yoshikazu Kuwana2), Hideaki Sugawara3)
1)Fujitsu Limited, 2)Tokai Software, 3)National Institute of Genetics

Contact us: sugawara@genes.nig.ac.jp

GBIF Japan portal site is supported by National BioResource Project (NBRP).

◆ Summary of GBIF
- In 2011, GBIF will provide 1 billion data of specimen or observation on 1.8 million species.
- GBIF is managed by the GBIF member countries.
- Biodiversity data is open to the public by organizations in the GBIF member countries or NGOs.
- Statistics as of Dec. 2009

<table>
<thead>
<tr>
<th>Data provider such as university or museum</th>
<th>303 organizations / 39 countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biodiversity information such as specimen or observation</td>
<td>190 million</td>
</tr>
</tbody>
</table>

◆ Roles of National Institute of Genetics (NIG)
- NIG collects domestic specimen/observation data and register to international GBIF database.
  - From 20 domestic organizations such as Tokyo university, national science of museum, Hyogo museum and so on.
  - 510 thousands of specimen/observation data such as mouse, insect, plant and Lact. bacillus are registered to GBIF from NIG.
- Promote the activity of biodiversity informatics in Japan.
- Annual symposium
- GBIF Japan Portal site

GBIF日本ポータルサイト

Acceptance of GBIF data

GBIF Japan portal site translate from Japanese name to GBIF from NIG.

GBIF search by Japanese common name

- Translation from scientific name to Japanese common name is done by a dictionary developed by Database Center for Life Science.
- GBIF headquarter does not accept Japanese common name.
- GBIF Japan Portal site translate from Japanese name to scientific name and search specimen/observation data.

Example of searching by Japanese “TUBAME (Barn Swallow)”

Distribution chart of TUBAME

◆ Examples of specimen or observation data
- Scientific name: Mus musculus Linneaus, 1758
- Taxonomic information:
  - Kingdom: Animalia
  - Phylum: Chordata
  - Common name in each country
    - House Mouse (English)
    - Souris Commune (French)
  - Year, month and day on which the specimen was collected.
  - Location of specimen collected or observation such as latitude, longitude, country, state and county.
  - Display of distribution of specimen or observation

Please refer to the under left distribution chart of Barn Swallow observation.

- Images of specimen or observation data

◆ Procedure of data registration to GBIF
1. To map your specimen/observation data to Darwin Core which is the GBIF data standard.
2. Examples of Darwin Core items (Most parameters are optional.)
   - Catalogue number (e.g. CI-119467)
   - Code of provider’s organization such as university or museum (e.g. HYO in the case of Hyogo Museum)
   - Record type (e.g. observation/specimen/living organism/germplasm/seed)
   - Scientific name and taxonomy (e.g. Kingdom, Phylum, Class, Order, Family, Genus, Species)
3. The GBIF Japan node checks the file as follows.
   - Whether the required items are filled in.
   - Whether values are valid. For example, latitude and longitude should be numeric.
   - Whether catalog number is unique within the data set.
4. To send the GBIF Japan node the data in an Excel or a tab delimited file which conform with Darwin Core.
5. The GBIF Japan node registers and configures the data to GBIF from NIG.

GBIF search box (Japanese common name is acceptable.)

http://gbif.ddbj.nig.ac.jp/

◆ Species Identifying System by Barcode
- Barcode of Life project is building a database of partial DNA sequences of Cytochrome Oxidas Subunit 1 (COXI) as barcodes for the identification of animal species.
- When two high school girls of New York collected fishes in the restaurant and analyzed them by Fish BoL database in a university, it turned out that 14 fishes of 56 were “Imitation” and it was reported by New York Times etc.
- Species Identifying System by using BoL data and 16S rRNA for bacteria is open to the public by GBIF Japan portal site.

Species Identifying System

Phylogenetic tree of identifying species