

BioResource Now!

Issue Number 7 June 2011

News from
Information
Center

Ongoing
Column No. 58

Hideaki Sugawara, National Institute of Genetics

Global Biodiversity Information Facility (GBIF) moves forward to the 3rd Phase <II>

P1 - 2

Use Windows 7 and IE9 for further convenience

P2

News from Information Center

Global Biodiversity Information Facility (GBIF) moves forward to the 3rd Phase <II>

Hideaki Sugawara

Specially Appointed Professor, National Institute of Genetics

What does GBIF Japan National Node do?

2) Administration of the GBIF portal site

The 2nd role of the Global Biodiversity Information Facility (GBIF) of the Japan National Node is to develop an information distribution system that is adapted to each country or region, the aim being to promote the use of biological diversity information. Through the GBIF Japan portal site (<http://www.gbif.nig.ac.jp/>), the National Institute of Genetics (NIG) supports the release of GBIF data, as introduced in article 1) of the prequel issue. The website has a GBIF search function that allows one to search in Japanese by common names. The website also has links to useful websites and provides introductory information about journal articles, events, international trends, and domestic activities regarding biodiversity. The page with useful links has tags to and brief descriptions of 247 websites.

The GBIF portal site also provides lecture handouts of "Symposium on Biodiversity in the 21st Century," which has been jointly hosted by NIG, the University of Tokyo, and the National Museum of Nature and Science (NMNS) every year since 2006.

Lecture Slides

- 2006: Creation of Biodiversity Informatics
- 2007: Creation of Biodiversity Informatics 2
- 2008: Development and Application of Large-Scale Information Networks Concerning Environment and Biodiversity
- 2009: Accumulation and utilization of biodiversity and observation in conjunction with distribution information of species
- 2010: Why Are the Scientific and Japanese Names of Species so Complicated? -A Key for Searching for Biodiversity Information

The GBIF Japan portal site also provides a species identification system (<http://bol.ddbj.nig.ac.jp/>) that uses the Barcode of Life Data Systems (<http://www.boldsystems.org/views/login.php>; Sugawara 2011^{※1}), which is affiliated to GBIF and the International Nucleotide Sequence Databases (INSD; <http://www.nsidc.org/>). As of May 2011, species can be identified on the basis of the following sequences: CO1, rbsL, ITS, matK, trnH-psbA, 16S rRNA, or gyrB.

3) From Data to Analysis on Distribution Information

(a) Distribution Information

Distribution of a species on a map can be observed by entering the Japanese name of the species in the search box on the GBIF Japan portal site (Fig. 2). For example, distribution information of Japanese knotweed can be obtained as shown in Fig. 2 if one inputs

the Japanese name of Japanese knotweed, "Itadori," as the keyword; selects "*Fallopia japonica*" from the candidate scientific names listed; selects "Species" from the expanded scientific name list containing "Subspecies," "Variety," etc.; and accepts the GBIF terms of use.

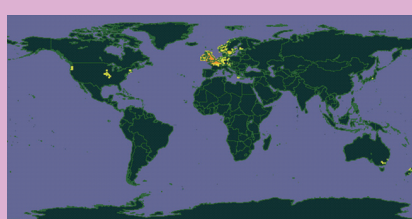


Fig. 2: Distribution of Japanese knotweed indigenous to Japan. The color in a cell surrounded by one degree of longitude and latitude becomes more red than yellow in proportion to the distribution density of a species. Regions colored green are those where the species is not found. The image is taken from the webpage on Japanese knotweed on the GBIF Japan portal site.

(b) Prediction of Biodistribution (Niche Modeling)

As a result of the collaboration between the GBIF system and the openModeller (<http://openmodeller.sourceforge.net/>), the bio-distribution of a species can be predicted in approximately 5 min with only 4 clicks from the distribution map screen described in (a), as long as the default parameters are used. The NIG prepared and released a video tutorial of niche modeling using distribution prediction of Japanese knotweed, indigenous to Japan, as an example (http://gbif.ddbj.nig.ac.jp/niche_model/). The conclusion of the video tutorial is shown in Fig. 3. By comparing Fig. 3 with Fig. 2, it is clear that the predicted distribution outside Japan matches observed results and includes regions in the Southern Hemisphere, where there are no records of this species being found.

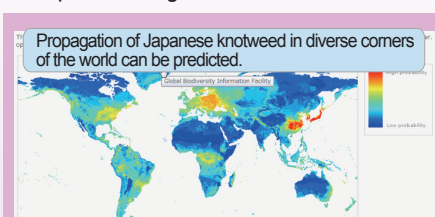


Fig. 3: Example result of biodistribution prediction (niche modeling). Worldwide propagation areas of Japanese knotweed are predicted based on observation data from Japan. The image was captured from the video tutorial.

(c) Crosstalk with Information Resources outside GBIF

Affected areas or areas potentially affected by an infectious disease are predicted by combining distribution information of a species recorded in GBIF, and information regarding the infectious disease that the species transmits (①-④).

① Extraction of a relationship between an infectious disease and its vector species, using Freebase (<http://www.freebase.com/>).

Freebase contains more than 20 million units of structured data. Knowledge about malaria, for example, is described by the Resource Description Framework (RDF; http://en.wikipedia.org/wiki/Resource_Description_Framework) in Freebase as follows.

- /en/Malaria (Subject) type (Predicate)
/medicine/infectious_disease (Object)
: Malaria is a type of infectious disease.
- /en/Malaria (Subject)
/medicine/infectious_disease/vector (Predicate)
/en/anopheles_gambiae (Object)
: Malaria is transmitted by the vector species *Anopheles gambiae*.

In this manner, by using structured data, one can obtain the relationship between an infectious disease and its vector species, which is not as easily extracted with conventional search engines.

② Expansion of information regarding an infectious disease and its vector species, using Animal Diversity Web.

Animal Diversity Web (<http://animaldiversity.ummz.umich.edu/site/index.html>) contains information on various species with literature citations. The website has a section titled "Economic Importance for Humans: Negative," which describes information on infectious diseases that are transmitted by vector species. We checked the information at NIG, extracted the relationship between the infectious disease and its vector species, and supplemented the information obtained from Freebase.

③ Expansion of Freebase

Anyone can create an account and register structured information on Freebase. We registered the relationship between the infectious disease and its vector species extracted from the Animal Diversity Web to Freebase.

■ Reference

- ※1 Hideaki Sugawara: "Biodiversity Information and the Barcode of Life" / *DNA Polymorphism* / Vol. 19:1-12, 2011 (Japanese only)

↳ To the next page

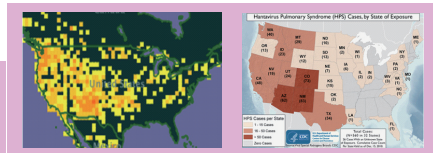
④ Investigation of Observation Information for a Vector Species

Since the relationship between an infectious disease and its vector species can be extracted from the expanded Freebase, we compared the information on *Peromyscus maniculatus* (deer mouse) distribution in the United States, which was obtained from GBIF, and areas affected by hantavirus infection in the United States. By cross-checking the right and left images in Fig. 4, it is clear that distribution patterns in the east and west are analogous. Therefore, if more accurate and updated data were available, it would be possible to devise preventive measures for an infectious disease, using the distribution information of its vector species.

Fig. 4: Distribution of *Peromyscus maniculatus* (deer mouse) (left) obtained from GBIF and areas affected by hantavirus infection (right).

Future plans

GBIF has existed for 10 years now. During this time, new information technologies have emerged, such as Linked Data (<http://linkeddata.org/>). Applications of these new information technologies provide an opportunity to develop more sophisticated distributed integrated systems. Nevertheless, the need of the moment is data accumulation. More exhaustive and reliable time-course data are required to address Access and Benefit Sharing (ABS), manage preventive measures for biodiversity, predict invasive species, or analyze distribution of an infectious disease. Even in the United States and United Kingdom, which are



first and second, respectively, in terms of amount of released GBIF data, only about 5% of the information from samples maintained in large-scale museums has been digitized (GBIF Strategic Plan 2012–2016). Although it is a meticulous task, digitization of data recorded on cards and notebooks and publication of the digitized data according to GBIF standards are of foremost importance to GBIF.

At the same time, although sample information is crucial for measuring biodiversity, it is necessary not only to accumulate observation data but also to promote cooperation with the Barcode of Life in order to accomplish the slogan of the 3rd phase of GBIF, "Seizing the Future to Benefit Science and Society." Further discussions regarding this matter will occur at the 18th meeting of the GBIF Governing Board organized in Argentina in June 2011.

"translated by ASL translation service, and validated by the author"

Use Windows 7 and IE9 for further convenience

Microsoft started the download distribution of the Japanese edition of Internet Explorer 9 (IE9) in April 2011. First, I will introduce the "pinning" feature, which can bookmark a user's favorite website to the task bar as an icon in the same manner as a general application.

In addition, I will also introduce how to customize the "Jump List," which can be displayed by right-clicking the icon pinned on the task bar and used to open an often-used menu for each site with one click.

1 Pinning Feature

The "pinning" feature is introduced first. In order to bookmark a website to the task bar, simply drag and drop either the icon on the left of the URL displayed or the tab of the website.



Fig. 1: Procedure to pin a website

Ongoing Column [No. 58]



Once a website is pinned, an icon of the website is added to the task bar, and the pinned website can be accessed by clicking the icon.

The icon of the website is displayed on the upper left of the window (Fig. 2).

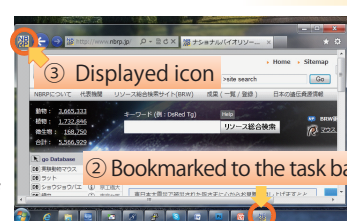


Fig. 2: Image of a pinned website

2 Jump List Feature

Next, a procedure to customize the "Jump List" is introduced.

In this example, I bookmarked the links ("About the NBRP," "NBRP Centers," "Resource Integrated Search Site (BRW)," etc.) on the upper menu bar of the public information website of the National BioResource Project (NBRP; <http://nbrp.jp>). If "About NBRP" is to be displayed, for example, the following meta tag should be added between <head> and </head> in the HTML code.

```
<meta name="msapplication-task"
content="name=About the NBRP;
action-uri=http://www.nbrp.jp/about/about.jsp;
icon-uri=http://www.nbrp.jp/image/favicon.ico" />
```

The format of the meta tag above is as follows.

```
<meta name="msapplication-task"
content=" (Name to be listed in the Jump List);
action-uri=[URL to be transferred upon clicking on the item in the Jump List];
icon-uri=[URL of the icon displayed on the left of the item in the Jump List]" />
```

※ Use of Internet Explorer 9 operated on a Windows PC is intended at this time.

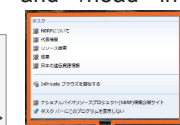


Fig. 3: Customized Jump List

(Gaku Kimura)

Reprinting and reduplication of any content of this newsletter is prohibited. All the contents are protected by the Japanese copyright law and international regulations.

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

The special topic on resources in the next month's issue will be "Yeast Genetic Resource Center."

Contact Address

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



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 7 June 2011

Editor's Note

Anyone can freely use GBIF's modeling feature, and a video tutorial is also available on the website of the GBIF Japan National Node. Please try to use the feature with a species of your interest. Doing so, could help you visualize biodiversity on a global scale and realize that searching with the Japanese name of a species can be convenient. (Y.Y.)

"translated by ASL translation service"