

BioResource Now!

Issue Number 9 December 2013

Hot News
No. 44

Tsuyoshi Hosoya (National Museum of Nature and Science) Localization and the Japan Node of the Global Biodiversity Information Facility (GBIF)

P1 - 2

Ongoing Column
No.85

Easy and Free Mapping of NGS Data to Genomic Data - Part II -

P2

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/>

Hot News <NO. 44>

Localization and the Japan Node of the Global Biodiversity Information Facility (GBIF)

Tsuyoshi Hosoya
Manager, Japan Node of GBIF,
National Museum of Nature and Science

Do You Know About the GBIF?

GBIF is well-known in the field of biodiversity science. However, one does not yet hear this conversation in Japan: "Do you know GBIF?" "Of course I know it. It is an abbreviation of the Global Biodiversity Information Facility." Well, the GBIF is an international organization, funded by governments, that through open data infrastructure provides information on biodiversity to anyone from anywhere in the world via the Internet. The GBIF was established in 2001 and has enjoyed the support of the National Institute of Genetics since its inception. Previous articles regarding the GBIF are available at

http://www.shigen.nig.ac.jp/shigen/news/n_letter/2006/nl200602En.pdf

http://www.shigen.nig.ac.jp/shigen/news/n_letter/2011/nl201105En.pdf

http://www.shigen.nig.ac.jp/shigen/news/n_letter/2011/nl201106En.pdf

In this article, I discuss the recent activities of the GBIF and its Japan Node.

Activities of the GBIF

The GBIF is funded by subscriptions from member countries, and non-governmental organizations that have signed the Memorandum of Understanding. With its headquarters in Copenhagen, Denmark, the GBIF collects specimens, observational information, and data with regard to biological resources. A biological specimen is assured evidence that an organism once existed in the area where the specimen was collected. Although the conservation of biological specimens has been practiced for a long time, information on biological specimens has not been sufficiently digitized and the considerable potential of databases of records of biological specimens has not been fully exploited. The GBIF has collected data from all over the world and, at present, more than 410 million species occurrence records are available through its website (<http://www.gbif.org/>).



Photo 1: Headquarters of the GBIF
The organization's headquarters are located on the second–fourth floor of the Zoological Museum at the University of Copenhagen.



Photo 2: Workstations in the headquarters
Approximately 20 full-time secretaries are employed at the organization's headquarters, with several secretaries sharing an open plan environment.

GBIF Participant Nodes in Asia

An increase in the number of participants has impacted the management of the GBIF. The secretariat of the GBIF divides its areas of activity into six regions: Asia, Africa, Europe, North America, Latin America, and Oceania. Each region functions independently. However, the GBIF Participant Nodes in Asia, to which the Japan Node belongs, have not been able to successfully integrate their activities due to regional diversity in language, culture, and natural environment, and so on.

However, since the election of Dr. Sheila Vergara of the ASEAN Center of Biodiversity in the Philippines as the new Asian region representative of the GBIF at its regional meeting in March 2012 in Tokyo, the GBIF Participant Nodes in Asia have made considerable progress towards integrating the regions' activities. In addition, strategies for the GBIF Participant Nodes in Asia were established at the subsequent meeting held in June 2012 in Taipei. Although the term "strategy" may sound broad, it clearly indicates a commitment to establishing a direction or framework aimed at achieving certain goals in the long-term. The Taipei meeting also determined that participating countries were to pursue the following activities based on the established strategies: (1) creation and integration of species checklists (particularly invasive alien species, endangered species, endemic species, and migratory birds); and (2) research on biodiversity loss and risk assessment in Asia using fish databases.



Photo 3: Taipei meeting
In addition to the two people from the secretariat, representatives from Taiwan, India, Japan, the Philippines, and Pakistan participated in the meeting.

<Strategies for the GBIF Participant Nodes in Asia>

- 1) Build a network of data holders.
- 2) Recommend data papers for publication.
- 3) Promote the use of specimen data that has not been used but is stored in museums and specimen depositories.
- 4) Foster the use of Darwin Core Archives as a standard for archiving data.

Strategies for the Japan Node of GBIF

As many participants in the GBIF are countries (some are research institutes), a node has been formed within each participating country. The Japan Node of GBIF is known as JBIF. Every 5 years, the GBIF adopts a slogan and establishes medium-term goals and associated strategy. The current slogan is: "seizing the future" and the strategic goals are:

- 1) Advance digital data;
- 2) Advance informatics infrastructure;
- 3) Advance engagement.

The primary role of the JBIF is to promote the use and global dissemination of biodiversity-related data obtained in Japan. The JBIF has established its strategic goals in light of the strategies for the Asian region and the entire GBIF.

<Strategic goals established by the JBIF>

- 1) Increase awareness of the importance of biodiversity information in researcher communities.
- 2) Improve the biodiversity-related function at museums.
- 3) Communicate the importance of biodiversity information to the public and policy makers.
- 4) Strengthen the Japan Node's organizational presence.
- 5) Seek partnerships with related projects.
- 6) Demonstrate leadership in collaborative activities within Asia.

Communicate More Actively

Both the GBIF and the JBIF are organizations that collect and publish data. To collect data effectively, the organizations' data collection systems must reflect an understanding of the role researchers play and must be managed collectively by a group of researchers. Thus, enhancing transparency with regard to both organizations' activities is essential to ensure the work of the organizations is effectively communicated to a wide group of people.

Easy and Free Mapping of NGS Data to Genomic Data - Part II -

Ongoing Column [No. 85]



Visualization using IGV

The previous edition introduced a genome-mapping tool called Bowtie2. This edition introduces the Interactive Genome Viewer (IGV), a tool used to view the output generated in the previous step.

1 Downloading the genome viewer

Click on the "Download" link using the URL below to download IGV_2.3.23.zip (the latest version as of December 20, 2013) and unzip its contents.

URL : <http://www.broadinstitute.org/software/igv/download>

If you are downloading IGV for the first time, you will have to register your e-mail address.

Java is required to run this tool – you can download and install the latest Java version from this URL:

URL : <http://java.com/en/download/>

2 Data preparation

The reference genome (in FASTA format) used for the map sequence and the *sam* file created in the previous step are used for visualization. You can optionally specify the gene data file (in *gff* format) that was downloaded with the reference genome.

3 Importing the reference sequence

To launch IGV, double-click on the "igv.jar" file included in the unzipped folder. The application comes with genome sequences of the major model organisms. However, we will import the reference sequence data used in the previous mapping step in this example. Click on [Genomes – Create Genome File] from the menu and enter an arbitrary name in the [Unique identifier] and [Descriptive name] fields. Click

on the [Browse] button next to the label "FASTA file" and select the reference sequence file (Brapa_sequence_v1.5.fa). In addition, select the gene file previously downloaded (Brapa_gene_1.5.gff) in the [Gene file] field as an option (Fig. 1). Click [OK] to create the genome file, which will then be loaded into IGV.

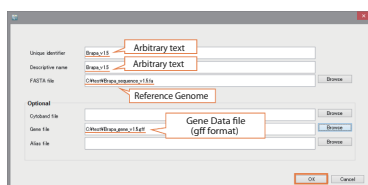


Fig.1. "Create .genome File" screen

4 Preparing the sam file

The *sam* file created using Bowtie2 is unsorted; it must be sorted before it can be used. Click on [Tools – Run igvtools] from the menu bar. Select "sort" from the [Command] pull-down menu and select the *sam* file (test.sam) created using Bowtie2 for the [Input File] field. Click [Run] to generate a sorted file (Fig. 2).



Fig. 2. Sorting a sam file

Once the sort has completed, we must create the index file. Change the [Command] to [Index], and select the sorted *sam* file (test.sorted.sam) in the [Input] field. Click [Run] to generate the index file (Fig. 3).

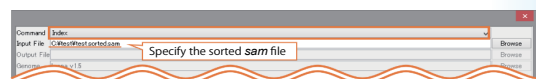


Fig. 3. Generating the index file

5 Loading the sam file

Click on [File – Load from File] on the menu bar and select the sorted *sam* file (test.sorted.sam) to load the mapped sequence data.

Once the loading has completed, select "Chromosome" from the menu located at the top and select a region in order to view the mapped sequence data. You can also jump to a specific coordinate by specifying the ID used in the *gff* file (Fig. 4).

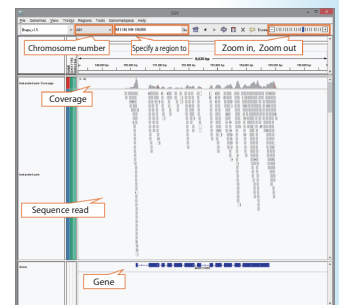


Fig. 4. Genome viewer

(Shunsuke Maeda)

Editor's Note

By 2011, the GBIF had collected 270 million records of published data. Thus, the amount of data collected has increased approximately 1.5 times in two years. Although funds for its activities are not always sufficient, the GBIF has achieved steady results. In 2012, Dr. Tsuyoshi Hosoya of the National Museum of Nature and Science was appointed as the manager of the JBIF. Since his appointment, he has been actively involved with various activities and has established a working group comprising young researchers. I am very proud of the Japanese team's considerable contribution to the work of GBIF in the Asian region.

This year, Dr. Kazuo Moriwaki, one of the founders of the National BioResource Project, passed away in November. We are grateful for his support of the Japan Node of GBIF as its Steering Committee Chairman in past year. May his soul rest in peace. I am committed to carrying on Dr. Moriwaki's work and contributing more next year. (Y. Y.).

Contact Address

Genetic Resource Center, National Institute of Genetics
1111 Yata, Mishima-shi, Shizuoka 411-8540, Japan
Tel.: 055-981-6885 (Yamazaki)
E-mail : brnews@shigen.info

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 9 December 2013

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