

BioResource Now !

Issue Number 9 October 2013

Hot News
No. 43

Masahiro Kishii (International Maize and Wheat Improvement Center)

CIMMYT Wheat Breeding and Japan's Contribution to World Wheat Production

Ongoing Column No.84 Easy and Free Mapping of NGS Data to Genomic Data

Database of This Month National BioResource Project "Zebrafish"

P1 - 2

P2

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.43>

CIMMYT Wheat Breeding and Japan's Contribution to World Wheat Production

Masahiro Kishii

Head of Global Wheat Program,
International Maize and Wheat Improvement Center

The International Maize and Wheat Improvement Center (CIMMYT; located in Mexico; **Photo 1**) is an international breeding center aiming to improve living standards of poor farmers in the world by providing better wheat and maize varieties. One of the achievements of the institute is a pioneer work on the development of semi-dwarf wheat varieties headed by Dr. Norman Borlaug that led to the Green Evolution. Since then, many of the wheat varieties produced in CIMMYT have been not only directly released to developing countries where their own breeding capacity has not been sufficient but also utilized as breeding materials in many industrialized countries.

As predicted by a report that the world population will increase to 10.9-16.6 billion by the end of this century (UN DESA 2013; <http://www.un.org/en/development/desa/index.html>), it is necessary to keep releasing higher yield wheat varieties to satisfy increasing food demand. But in recent years it has become more difficult to catch up this mandate because of global climate change that would cause many abiotic stresses such as drought and heat.

To overcome these difficulties, CIMMYT wheat-breeding program has been utilizing any kind of key methodologies that would be able to facilitate for releasing higher yield and stress resistance varieties, including molecular technique, bio-informatics (such as genomic selection), transgenic, physiological approaches, F1 hybrid wheat, improved agronomical practices and, of course, use of genetic resources (landraces, wheat ancestor and alien species).

It is not known by many Japanese people that Japan has been hugely helping the works of CIMMYT by providing infrastructures including its Genebank building (**Photo 2**). Much less people are aware of the more important fact that hundreds of accessions of wheat ancestor species (namely *Aegilops tauschii*) originally coming from NBRP-Wheat (or Kyoto University prior to the start of NBRP project) have been used to produce new synthetic bread wheats (man-made new bread wheat by crossing durum and *Aegilops tauschii*; **Photo. 3**)



Photo: Dr. Kishii in front of the Genebank Building



Photo 1: CIMMYT headquarter in Mexico



Photo 2 : The front of the Genebank building in CIMMYT



Photo 3: New synthetic wheat (right; green color as resistance for rust disease) and non-resistant wheat variety (left)

↳ To the next page

for improvement of drought and other stresses (Photo. 4). It is certain that many of additional genetic resources in NBRP-Wheat will contribute to world wheat production in the future.



"This article is written in English by author."



Photo 4 : Drought tolerance test in Obregon field. The right side is for drought condition, and the left is for control with irrigation.

Easy and Free Mapping of NGS Data to Genomic Data – Part 1 –

Advances in next-generation sequencing (NGS) methods and their wide availability have led to a greater number of sequence databases, making it easier to acquire NGS data in recent years. I will introduce techniques for mapping NGS data to genome sequences and viewing the results using a genome viewer. The article is split into two parts, with Part 1 being presented this month.

※ Although examples in this article use Windows (validated in Windows 7 and 8), the software being introduced also supports Mac and Linux operating systems.

Mapping sequence data to genomes

A tool called Bowtie2 will be used for mapping sequence data.



In the example below, a folder called "test" is created directly under the C:\ drive. Downloaded zip files as well as unzipped files are stored in this folder. Please use an appropriate tool for extracting the contents of the downloaded archives.

1 Obtaining the software

You can download Bowtie2 from the following URL :

<http://sourceforge.net/projects/bowtie-bio/files/bowtie2/2.1.0/>

Download bowtie2-2.1.0-mingw-winXX.zip (XX= 64 for 64-bit OS or 32 for 32-bit OS) and extract the contents to a folder.

Open [Control Panel → System → Advanced System Settings] and click on [Environmental Variables]. Click on the "Path" variable within [System variables] and click on [Edit]. Add a semicolon (;) at the end of the input field and add the path to the unzipped content (Fig. 1). Please do not make changes to any of the existing path information.

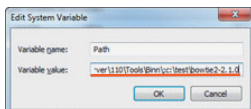


Fig. 1: Editing system variables

2 Preparing the data

You will need a reference genome (in FASTA format) and sequence data (in FASTQ or FASTA format). In the example below, we will use the genome data of *Brassica rapa*. This can be obtained from URL : <http://brassicadb.org/brad/ftpTrans.php> by clicking on v1.5/ and downloading the file, "Brapa_sequence_v1.5.fasta".

The sequence data can be obtained from the DNA Data Bank of Japan (DDBJ): access DRASearch (<http://trace.ddbj.nig.ac.jp/DRASearch/>), enter "ERX014878" in the "Accession" field, and click on [Search]. Click on the "FASTQ" link at the top of the page, which will take you to an FTP site. Download the file "ERR037339.fastq.bz2" and extract its contents. Place the extracted fastq file in the test folder created previously.

3 Mapping using Bowtie2

Click on the [Start] button, enter "cmd" in the search box, and execute "cmd.exe" or "command prompt." Type the command given in Table 1-1 and press [Enter]. Next, execute the command given in Table 1-2 in order to create an index. After the indexing has completed, execute the command given in Table 1-3 to commence mapping. The results will be displayed after a while and a file in SAM format will be generated. A SAM-format file comprises a header section starting with "@" followed by a standard set of fields.

Table 1. Commands (; : a space)

Index	Purpose	Command	Example
1	Move to a folder	cd _/d_(path to folder)	cd _/d_c:¥test
2	Create an index	bowtie2-build.exe _ {reference sequence file} _{index name}	bowtie2-build.exe _ Brapa_sequence_v1.5.fasta _ brapa_genome
3	Execute mapping and output a file in	bowtie2-align.exe _-x _ {index} _-U _{sequence file} _-S _{output}	bowtie2-align.exe _-x _ brapa_genome _-U _ ERR037339.fastq _-S _ test.sam

The generated SAM file is very large: please use a viewer that can handle large files. The standard fields in the SAM file are given below in Table 2.

Table 2. The contents of a SAM file

Row 1	Read name
Row 2	Read status (determines how the mapping is made)
Row 3	Chromosome or contig name
Row 4	Mapping position
Row 5	Mapping quality
Row 6	Mapping status (determines indel and matching rate)
Row 7	Name of mate in case of paired end
Row 8	Position of mate in case of paired end
Row 9	Insert length in case of paired end
Row 10	Read sequence
Row 11	Read quality

In Part 2 of this article, I will introduce Integrative Genome Viewer (IGV), which can be used to visualize the mapping results contained in the generated SAM file. (Shunsuke Maeda, Genetic Resource Center)

Database of This Month

National BioResource Project "Zebrafish"



• Number of strains: 357 (as of October 2013)

DB name : NBRP Zebrafish
 URL : <http://www.shigen.nig.ac.jp/zebra/>
 Language : Japanese, English
 Contents :
 • Zebrafish strain resources for research (wild, mutant, and transgenic strains)
 • Useful information on zebrafish (zebrafish-related sites, protocols, seminars, laboratories)
 • Newsletter
 Features : You can order resources from this database.
 Direct link with the Zebrafish Model Organism Database (ZFIM) enables you to obtain zebrafish-related information.
 Cooperative DB : ZFIN, zTRAP
 DB construction group : NBRP Zebrafish, NBRP Information Management organization : Genetic Resource Center, NIG
 Year of DB publication : 2004 Year of last DB update: 2013

Comment from a practicing developer: Although the NBRP Zebrafish is a small and simple database in terms of the amount of data and information items, the database can provide sufficient information to users, in cooperation with ZFIN. The database is managed by the Center for Genetic Resource Information, National Institute of Genetics. Since information on zebrafish is managed by NBRP Zebrafish, the latest information can be provided. In the future, we will add research papers from users of our resources to the database. Please feel free to use the database, and do not hesitate to send us your comments, questions, or opinions via Contact Us on the left-hand side menu.

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

Editor's Note

Dr. Kishii, who had temporarily returned to Japan to participate in the 12th International Wheat Genetics Symposium held in Yokohama last month, willingly agreed to write an introductory article on CIMMYT. I was surprised by the large-scale field for tolerance tests in CIMMYT, and I was pleased to know that Japan's wheat resources have greatly contributed to world wheat production. In April 2013, Dr. Kishii began his second stay at CIMMYT. I believe that Dr. Kishii will play an important role in CIMMYT (Y. Y.).

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 October 2013

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