

## Short Communication

# NIASGBsnp: integration of single nucleotide polymorphism data of rice (*Oryza sativa* L.) genetic resources

Masaru Takeya\*, Fukuhiro Yamasaki, Sachiko Hattori and Kaworu Ebana  
*Genetic Resources Center, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan*

Received 7 November 2012; Accepted 17 February 2013 – First published online 21 March 2013

### Abstract

The NIASGBsnp system manages data on single nucleotide polymorphisms (SNPs) of rice (*Oryza sativa* L.) genetic resources in the National Institute of Agrobiological Science (NIAS) Genebank. NIASGBsnp currently holds data on 768 SNP markers for 301 rice accessions and plans to add the SNP data of active rice accessions in the NIAS Genebank. It can show differences between accessions by graphical genotyping. Passport, characteristics and evaluation data of accessions can be retrieved to allow phenotype to be associated with genotype. NIASGBsnp will support various research purposes such as genomic selection and plant pathology research.

**Keywords:** database; genebanks; graphical genotype; *Oryza sativa* L; SNP; Web-based retrieval system

### Experimental

The National Institute of Agrobiological Sciences (NIAS) supports the NIAS Genebank Project for the conservation of plant, microorganism and animal genetic resources related to food and agriculture in Japan (Okuno *et al.*, 2005). Genetic resources are classified, evaluated, multiplied and preserved in the NIAS Genebank. Genetic resources in the public domain are distributed for research and educational purposes. The database holds passport, characteristics, evaluation and storage control data (Takeya *et al.*, 2011).

Rice (*Oryza sativa* L.) is an important crop and a model plant (e.g. International Rice Genome Sequencing Project, 2005). The NIAS Genebank currently has 20,079 rice (*O. sativa* L.) accessions from 128 countries that are

accessible. Of the conserved rice accessions, 88% have evaluation data. To enrich the genomic information available, the database now holds data on single nucleotide polymorphisms (SNPs). We have developed a system, called NIASGBsnp, to manage the SNP data (Fig. 1). Although several open rice SNP databases exist, such as the OryzaSNP Database (<http://www.oryzasnp.org/>) and the GRAMENE SNP Query ([http://www.gramene.org/db/diversity/snp\\_query](http://www.gramene.org/db/diversity/snp_query)), the linkage between the SNP data and pertinent accession information such as phenotypic data is restrictive in most cases. The acquisition of SNP data by genotyping bulk DNA samples from each accession supports the full integration of NIASGBsnp into the Web-based plant accessions search software. Associated diseases can be matched, and the pertinent set of host and pathogen can be accessed from the NIAS Genebank.

The genome-wide SNPs of 140 Asian rice accessions were surveyed to reveal the sequence diversity and population structure of the cultivars (Ebana *et al.*, 2010). Bulk total DNA was extracted from young leaves of

---

\*Corresponding author. E-mail: [katu@affrc.go.jp](mailto:katu@affrc.go.jp)

NIAS Genebank  
農業生物資源データベース

### Plant Search (Evaluation Data Queries)

For more information on how to search, please refer to the [Explanation of the Plant Genetic Resources Search System](#).

52 accessions found 3 / 3  
SNP data (52 accessions) are available.

| JP No. | Scientific name | Cultivar       | Status   | Origin | Availability | Standard Quantity (g) | Germination percentage |
|--------|-----------------|----------------|----------|--------|--------------|-----------------------|------------------------|
| 80529  | Oryza sativa    | MACK KHEUA HAY | Landrace | LAOS   | Yes          | 7.00                  | 88%                    |
| 12189  | Oryza sativa    | TEIKYAKU USEN  | Landrace | TAIWAN | Yes          | 7.00                  | 84%                    |

Search criteria

Evaluation group: Rice (Descriptors Evaluation data list)

Plant name:  Paddy rice  Upland rice  Wild rice  African rice

JP No.: JP: 12189

Conservation Lot No.:

Cultivar name:

Scientific name:

Origin: TAIWAN

Status: Landrace

Collection source: Introduced

Received date: 1967

Source:  Collected  Introduced  Bred  Derived  Unknown

Culm length (cm): 15 to 110

Panicle length (cm):

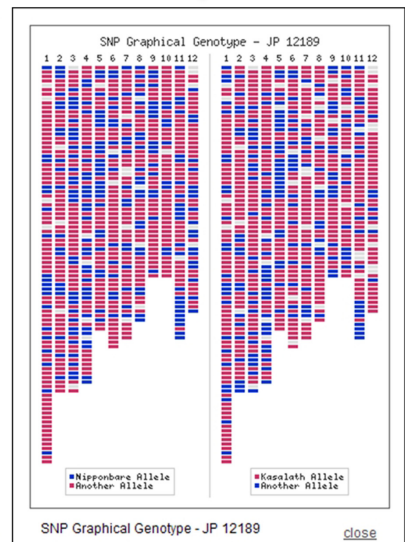
Number of panicles (Number per plant):

Apiculus color:  Straw  Tawny  Brown  Red brown  Light red  Red  Light purple  Purple  Blackish purple

Web-based plant accessions search software

| Acc | A                   | B | C        | D | E | F | G | H | I | J | K | L | M | N | O |
|-----|---------------------|---|----------|---|---|---|---|---|---|---|---|---|---|---|---|
| 2   | NIAS_Os_AHD1000003  | 1 | 25254    | A | T | A | A | A | A | A | A | A | A | A | A |
| 3   | NIAS_Os_AHD1000001  | 1 | 32667    | A | G | G | G | T | T | T | T | T | T | T | T |
| 4   | NIAS_Os_AHD1000058  | 1 | 503555   | G | A | A | A | G | G | G | G | G | G | G | G |
| 5   | PO175_3             | 1 | 520357   | G | G | A | A | G | G | G | G | G | A | G | G |
| 6   | NIAS_Os_AHD1000284  | 1 | 992163   | T | T | T | T | T | T | T | T | T | T | T | T |
| 7   | NIAS_Os_AHD1000314  | 1 | 1548551  | A | G | G | G | G | G | G | G | G | G | G | G |
| 8   | NIAS_Os_AHD10000510 | 1 | 1948598  | A | T | T | T | N | T | T | T | T | T | T | T |
| 9   | NIAS_Os_AHD1000769  | 1 | 2415712  | G | A | T | T | A | N | N | N | N | N | N | N |
| 10  | PO296               | 1 | 2718679  | C | C | T | T | T | T | T | T | T | T | T | T |
| 11  | NIAS_Os_AHD1000153  | 1 | 3013737  | A | G | G | G | G | G | G | G | G | G | G | G |
| 12  | NIAS_Os_AHD10001410 | 1 | 3390029  | A | A | G | G | G | G | G | G | G | N | G | G |
| 13  | PO182               | 1 | 4033610  | G | A | A | A | G | G | G | G | G | G | G | G |
| 14  | PO626_1             | 1 | 4771790  | G | T | T | T | G | G | G | G | G | G | G | G |
| 15  | NIAS_Os_AE01002154  | 1 | 5442929  | T | C | C | C | C | C | C | C | C | C | C | C |
| 16  | NIAS_Os_AHD1002655  | 1 | 5733591  | G | T | T | T | T | T | T | T | T | T | T | T |
| 17  | NIAS_Os_AHD1000396  | 1 | 6186197  | C | T | T | T | T | T | T | T | T | T | T | T |
| 18  | NIAS_Os_AHD1000437  | 1 | 6584409  | G | A | A | A | G | G | G | G | G | G | G | G |
| 19  | NIAS_Os_AHD1000472  | 1 | 6968931  | G | A | A | A | N | G | G | G | G | G | G | G |
| 20  | NIAS_Os_AHD1003587  | 1 | 7435035  | A | G | G | A | A | A | A | A | A | A | A | A |
| 21  | NIAS_Os_AE01002285  | 1 | 7813575  | A | C | C | C | C | C | C | C | C | C | C | C |
| 22  | NIAS_Os_AHD1000590  | 1 | 8230139  | C | A | C | C | C | C | C | C | C | C | C | C |
| 23  | NIAS_Os_AHD1000611  | 1 | 8627871  | G | T | G | G | G | G | G | G | G | G | G | G |
| 24  | NIAS_Os_AHD1000643  | 1 | 9178553  | A | G | G | A | A | A | A | A | A | A | A | A |
| 25  | PO631_1             | 1 | 9470048  | T | C | C | C | T | T | T | T | T | T | T | T |
| 26  | NIAS_Os_AHD10004243 | 1 | 10104516 | A | T | T | T | A | G | G | G | G | G | G | G |
| 27  | NIAS_Os_AHD1000747  | 1 | 10758269 | A | T | T | T | A | A | A | A | A | A | A | A |
| 28  | NIAS_Os_AHD10005077 | 1 | 11527288 | A | G | G | G | N | N | N | N | N | N | N | N |
| 29  | NIAS_Os_AHD1000885  | 1 | 12174422 | T | A | T | T | T | T | T | T | T | T | T | T |
| 30  | NIAS_Os_AHD1000966  | 1 | 12522150 | A | G | A | A | A | A | A | A | A | A | A | A |
| 31  | NIAS_Os_AHD1006120  | 1 | 13893938 | A | G | G | A | A | A | A | A | A | A | A | A |
| 32  | NIAS_Os_AHD1001076  | 1 | 14265892 | C | T | T | T | C | C | C | C | C | C | C | C |
| 33  | NIAS_Os_AHD1001131  | 1 | 14568427 | C | T | T | T | C | C | C | C | C | C | C | C |

SNP data spreadsheet



SNP Graphical Genotype - JP 12189

Fig. 1. Integration of SNP data server with Web-based plant accessions search software.

ten plants in each accession. A total of 4357 SNPs were identified by sequencing the exons and introns of anonymous rice genes on 12 chromosomes. Seven cultivar groups, including three tropical *japonica* and three *indica*, were identified by classifying the 140 accessions on the basis of these SNPs (Ebana *et al.*, 2010). From the 4357 SNPs, 768 SNPs were selected to characterize the rice accessions. We started to characterize all the rice active accessions in the NIAS Genebank. For the genotype characterization of the accessions, 25 plants per accession were used for DNA extraction to check heterogeneity in the accessions. SNPs were detected using the Golden Gate technology (Illumina, Inc., San Diego, CA, USA) with BeadStation 500G. To date, data on the 768 markers for 301 accessions have been registered in the genetic resources database.

The NIAS Genebank already held a lot of passport, characteristics, evaluation and storage control data with regard to many accessions with SNP data. The addition of more SNP data allows users to associate genotype with phenotype.

We have improved the Genebank search system for the return of SNP data. Under 'Plant Search (Simple Queries)' ([http://www.gene.affrc.go.jp/databases-plant\\_search\\_en.php](http://www.gene.affrc.go.jp/databases-plant_search_en.php)), target accessions are selected according to search criteria such as cultivar name, origin and collection source. Users can choose 'Accessions with SNP data only'. When such accessions are included in the search result, a link to download an XML spreadsheet file of the SNP data is presented above the results list. The spreadsheet lists the SNPs of the selected accessions and three references (Nipponbare, Kasalath and Dee-Geo-Woo-Gen). We developed the software to create an XML spreadsheet from the SNP data in the database.

Under 'Plant Search (Evaluation Data Queries)' ([http://www.gene.affrc.go.jp/databases-plant\\_search\\_char\\_en.php?type=1](http://www.gene.affrc.go.jp/databases-plant_search_char_en.php?type=1)), target accessions can be selected according to morphological characteristics, resistance to stresses, yields and a variety of other criteria. The option of 'Accessions with SNP data only' is also provided. An MS Excel spreadsheet listing all evaluation data can be downloaded from an XLS icon above the search results list. Datasets incorporating SNP, characteristics and evaluation data can be used in genomic selection or marker-assisted selection to reveal the associations between genotype and phenotype.

Detailed information on each accession can be displayed by clicking the accession ID ('JP No.'). SNPs can be displayed as a graphical genotype in the detailed information window (Fig. 1). The presentation of SNPs as a graphical genotype supports the intuitive understanding of users (e.g. Milne *et al.*, 2010). We developed a program to generate SNP images from the data in the database. The latest image is always available.

The reference accessions are Nipponbare (*japonica*) and Kasalath (*indica*). The markers are presented in the linear order of physical position on each chromosome. When the target accession is compared with Nipponbare, homozygous and heterozygous alleles are shown in blue and red, respectively. In contrast, when the target accession is compared with Kasalath, homozygous and heterozygous alleles are shown in red and blue, respectively. Thus, overall, *japonica* alleles appear blue and *indica* alleles appear red.

The NIAS Genebank database contains microorganism genetic resources that links to the Database of Plant Diseases in Japan ([http://www.gene.affrc.go.jp/databases-micro\\_pl\\_diseases\\_en.php](http://www.gene.affrc.go.jp/databases-micro_pl_diseases_en.php)) (Takeya *et al.*, 2011, 2012). The most common hosts in the database are 'rice' and 'rice grain'. In the window that opens when the corresponding disease name is selected, the database is linked to the plant and microorganism accessions search software via the 'Related hosts' and 'Related strains' link, respectively. The search criteria offer the new option of 'Having related hosts only' and 'Having related strains only'. The NIAS Genebank can supply pathogen and plant samples for plant pathology research. As the genetic resources database has increased genomic information such as DNA sequence data of the barcode gene regions of microorganisms (Takeya *et al.*, 2012), pathogen-related genotypic data will support various research purposes.

## Discussion

The SNP data server is integrated with the Web-based plant accessions search software of the NIAS Genebank and efficiently provides SNP data with passport, characteristics and evaluation data. Markers and accessions will continue to be added to the genetic resources database. The combination of genome-wide SNPs, polymorphisms of useful genes, phenotypic data and passport data will provide a powerful tool for selecting the accessions. It is possible to add a system to download flanking sequence data and a SNP query tool to find polymorphic SNPs between any two selected accessions. SNP data of other crops will be incorporated.

A unified database of microorganisms and plants including both of characteristics and genotype will enhance the research and breeding.

## Acknowledgements

We would like to thank our colleagues within the Genetic Resources Center at NIAS for their input into this work.

## References

- Ebana K, Yonemaru J, Fukuoka S, Iwata H, Kanamori H, Namiki N, Nagasaki H and Yano M (2010) Genetic structure revealed by a whole-genome single-nucleotide polymorphism survey of diverse accessions of cultivated Asian rice (*Oryza sativa* L.). *Breeding Science* 60: 390–397.
- International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature* 436: 793–800.
- Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas WTB, Flavell AJ and Marshall D (2010) Flapjack – graphical genotype visualization. *Bioinformatics* 26: 3133–3134.
- Okuno K, Shirata K, Niino T and Kawase M (2005) Plant genetic resources in Japan: platforms and destinations to conserve and utilize plant genetic diversity. *Japan Agricultural Research Quarterly* 39: 231–237.
- Takeya M, Yamasaki F, Uzuhashi S, Aoki T, Sawada H, Nagai T, Tomioka K, Tomooka N, Sato T and Kawase M (2011) NIASGBdb: NIAS Genebank databases for genetic resources and plant disease information. *Nucleic Acids Research* 39: D1108–D1113.
- Takeya M, Yamasaki F, Uzuhashi S, Kumagai M, Sawada H, Nagai T, Tomioka K, Sato T, Aoki T and Kawase M (2012) Development of a database of plant diseases in Japan and a system for making microorganism genetic resources and their DNA sequence data available to the research community. *Japan Agricultural Research Quarterly* 46: 193–198.