MARDI rice genebank: important roles in data management and data sharing

S Noorzuraini[1]*, M Shukri[2], A Amron[2], M Izzat[2], M Ramdzan[1] and N Idayu[1]

[1] MARDI Rice Genebank, MARDI Seberang Perai, Jalan Paya Keladi, 13200 Kepala Batas, Pulau Pinang, Malaysia
[2] My Gene Bank, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor, Malaysia

*E-mail: zuraini@mardi.gov.my

Abstract. Malaysia has agreed to adopt the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) and acceded to the Treaty in 2004. Under Article 10 of the Treaty, the contracting parties should allow easy access of the genetic resources under Annex 1 Crop through Multilateral System, and Benefit-Sharing raised from the utilization of the resources. To implement the Treaty, MARDI has actively involved in the implementation through the exchange of seeds particularly rice accessions using Standard Materials Transfer Agreement (SMTA) and in data sharing for easy access of rice genetic resources. MARDI has started to involve in adoption and implementation of Digital Object Identifiers (DOIs) and also contributed to the Genesys database though a project named 'Building Genesys Catalog of Phenotypic Datasets of Malaysian Rice Germplasm'. In Malaysia, rice genetic resources are conserved at the National Rice Genebank in Penang. The genebank was established in 1989 with the primary purpose to collect and conserve rice genetic resources (*Oryza* spp.) in Malaysia. Currently, the genebank has conserved a total of 13,020 accessions of rice germplasms consisted of landraces collected throughout the country, introduced varieties, and breeding materials. The rice germplasms are an important source of genetic materials for the development of new rice varieties. To date, 48 rice varieties were developed and introduced to the farmers and local communities. Those varieties possessed consequential and unique traits such as high yield and resistant to pest and diseases, to make them valuable to farmers and local communities. Those varieties are registered and freely accessed in Global Information System (GLIS) website for DOI implementation. Besides, a total of 22 phenotypic datasets including datasets for morphological traits, pest and diseases (brown planthopper, blast and bacterial leaf blight), and quality and speciality traits, are prepared and published in the Genesys portal (database). Data sharing is essential to enhance the breeding program in the country and sharing of valuable genetic resources with other countries for the future development of new varieties for food security and to promote sustainable agriculture.

Keywords: Malaysia, MARDI, DOI, rice genetic resources, ITPGRFA.

1. Introduction

Malaysia is located between 2° and 7° north of the equator and longitudes 100° and 119° east. This South-East Asian sovereign covers an area of about 329,758 km^2, consisting of Peninsular Malaysia, and the states of Sabah and Sarawak. Malaysia has been recognized as one of the world's twelve mega

diversity developing countries in the world. It harbours at more than 170,000 species [1] with more than 15,000 species of flowering plants [2]. The plant species are presented by more than 2,500 tree species, 3,000 species of orchids, 500 species of ferns, 60 species of grasses and bamboos, and many others [2].

Malaysia implemented the National Policy on Biological Diversity in 1998 and ratified the Convention on Biological Diversity in 1994. Ten years later, Malaysia agreed to adopt the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) and acceded to the Treaty in 2004. Under Article 10 of the Treaty, the contracting parties should allow easy access of the genetic resources under Annex 1 Crop through Multilateral System and Benefit-Sharing raised from the utilization of the resources. Consequently, MARDI has actively involved in the Treaty through the exchange of rice genetic resources by using the Standard Material Transfer Agreement (SMTA) and also in data sharing for easy access to the genetic resources. Recently, MARDI has involved in adoption and implementation of Digital Object Identifiers (DOIs) for rice germplasm and also contributed to Genesys portal (database) through a project named 'Building Genesys Catalog of Phenotypic Datasets of Malaysian Rice Germplasm' funded by Global Crop Diversity Trust (Crop Trust).

In Malaysia, rice genetic resources are conserved in the National Rice Genebank located at MARDI Seberang Perai, Pulau Pinang. The genebank was established in 1989 with the primary purpose to collect and conserve the rice genetic resources (*Oryza* spp.) in Malaysia. Currently, the genebank has conserved a total of 13,020 accessions of rice germplasm consisted of landraces collected throughout the country, introduced varieties, and breeding materials. These rice germplasms are important as genetic resources for research and development of new rice varieties in Malaysia. To date, 48 rice varieties were successfully developed by MARDI Rice Breeders and were introduced to the farmers and local communities. However, only 45 MARDI varieties were selected for DOI registration. These varieties possessed consequential and unique traits such as high-yielding, aromatic and colored, and resistant to pest and diseases, thus making them valuable to the farmers and local communities.

This paper will discuss on the establishment of data management in MARDI and the contribution of MARDI in data-sharing mechanism to support easy access of rice genetic resources established under the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA or the 'Treaty').

2. Data management in MARDI

To support the efforts for sustainable use, development, and conservation of plant genetic resources in the country, MARDI has developed a database information system known as MARDI Agrobiodiversity Information System (AgrobIS) (Figure 1). The information system was established to conserve data on collection and conservation of genetic resources related to agrobiodiversity components, namely plants, arthropods and microbes [3]. It is a web-based application and capable of handling specific characters of each agrobiodiversity component [3]. The system is an innovative technology for handling agrobiodiversity information using licensed free software, MySQL for database development and PHP for the web interface.

For rice germplasm, the information system consists of five categories: (1) Passport data–basic information of the accession such as accession number, variety name, original country, collector name, acquisition date and seed source; (2) Characteristics–consists of morphological characterization data on quantitative and qualitative traits. It is based on Rice Descriptor List developed by IRRI and Biodiversity International; (3) Evaluation–consists of screening data for biotic (drought, salinity) and abiotic stresses (brown planthopper, foliar blast and bacterial leaf blight); (4) Quality traits–consists of data on speciality traits of grain such as amylose content, gelatinization temperature, seed coat color, endosperm type and scent (aroma); (5) Seed Management–consists of information on management of the collection in genebank, i.e. specific location in medium and long term storage, duplicated area, regeneration details, seed quantity and result of germination test.
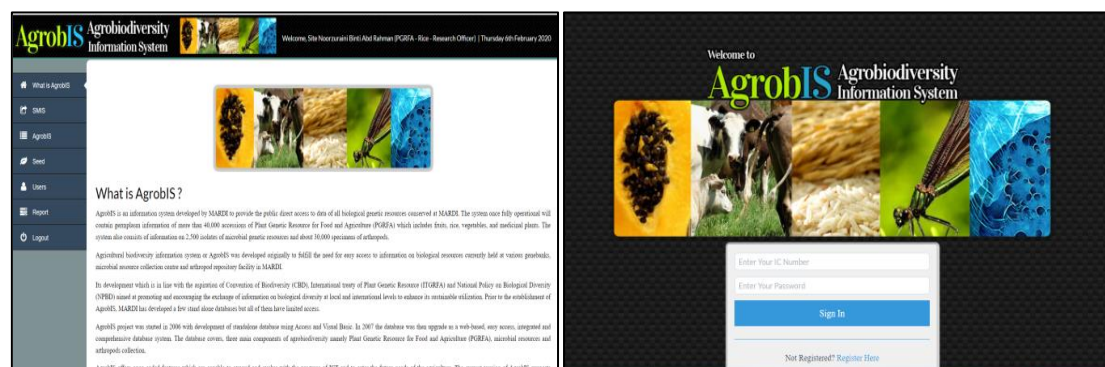
Figure 1. The overview of the Agrobiodiversity Information System (AgrobIS).

3. Building Genesys catalog of phenotypic datasets of Malaysian rice germplasm

MARDI has involved in 'Building Genesys Catalog of Phenotypic Datasets for Malaysian Rice Germplasm' as part of a larger Crop Trust project named 'A Genesys Catalog of Phenotypic Datasets Linked to Genebank Accessions'. The project started in 2017 and will end in August 2018. The objective of the project is to contribute to the building of metadata-based catalogue of accession-related phenotypic datasets on Genesys. In this project, 22 datasets of rice germplasm based on selected planting season were successfully published on Genesys (Figure 2). The datasets covered 624 rice accessions of landraces and introduced varieties for morphological characterization data (9 datasets), bacterial leaf blight (10 datasets), brown planthopper (1 dataset), foliar blast (1 dataset) and rice specialty traits (1 dataset).

The datasets for morphological characterization provide a broad spectrum of variation of morphological traits among the accessions (Table 1) which are fundamental in order to provide information for breeding program [4]. The highest coefficient of variation observed in culm number (33.11) followed with ligule length (31.98). Besides that, there are several accessions showed resistant for certain diseases. The bacterial leaf blight dataset observed 53 accessionsare resistant and 84 accessions are moderately resistant (Figure 3). Meanwhile, the foliar blast dataset observed nine accessions are highly resistant and 14 accessions are resistant to foliar blast disease (Figure 4). The rice specialty traits observed 25 accessions with low amylose content, 10 accessions with intermediate amylose content and five accessions with high amylose content (Figure 4). In Malaysia, most of the people preferred rice varieties with low to intermediate amylose content. The rice is less sticky and soft texture.
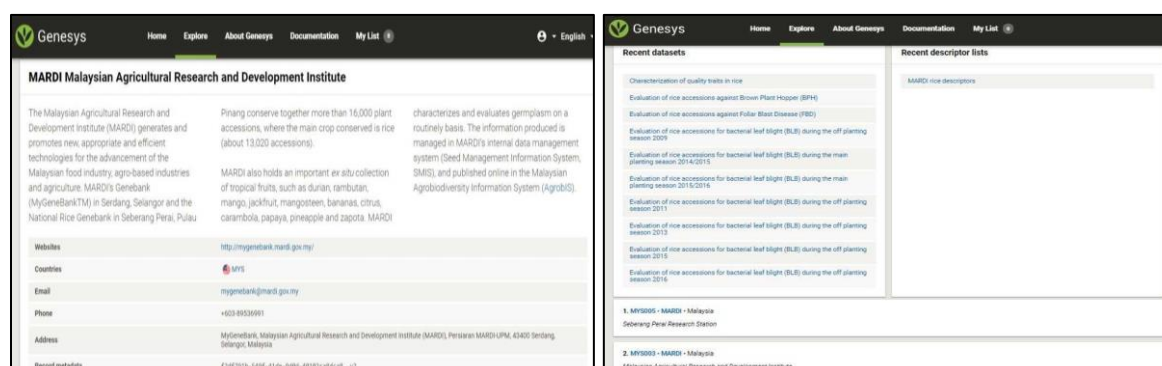

Figure 2. List of data on characterization and evaluation of Malaysian rice germplasm in the Genesys Catalog.

Table 1. The statistical analysis of selected rice germplasm in the Genesys Catalog.

Variable	N	Mean	Max	Min	Std dev	Std error	Variance	Coefficient of variation
Flag leaf length (cm)	153	34.87	73.6	15.0	9.06	0.73	82.06	25.98
Flag leaf width (mm)	153	12.83	21.8	6.2	2.62	0.21	6.84	20.38
Leaf length (cm)	143	44.65	82.6	16.2	9.95	0.83	98.95	22.28
Leaf width (mm)	143	10.81	18.8	6.0	2.21	0.18	4.88	20.44
Ligule length (mm)	142	16.27	38.7	5.2	5.20	0.44	27.08	31.98
Culm length (cm)	148	79.80	144.6	32.4	23.48	1.93	551.44	29.43
Culm number	148	14.29	32.0	5.0	4.73	0.39	22.37	33.11
Panicle length (cm)	149	24.05	34.8	9.8	3.77	0.31	14.22	15.68
Maturity	68	131.24	172.0	105.0	17.20	2.09	296.00	13.11
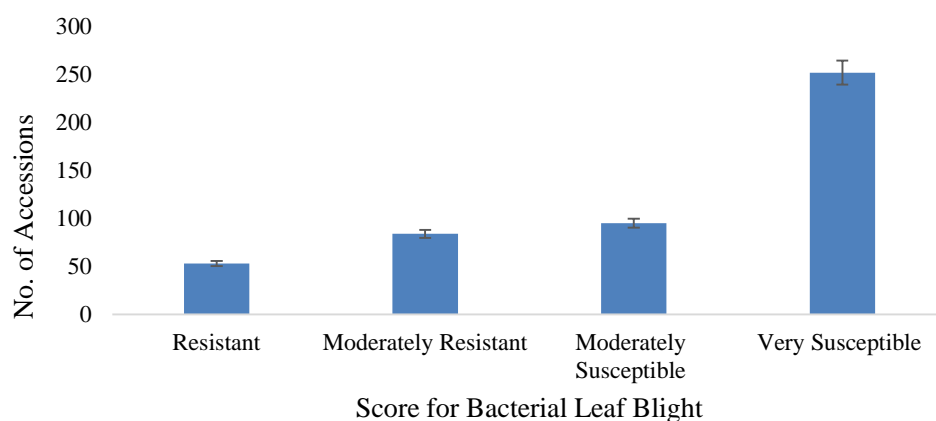
Figure 3. The score of bacterial leaf blight among the selected rice germplasm which are available in the Genesys Catalog.
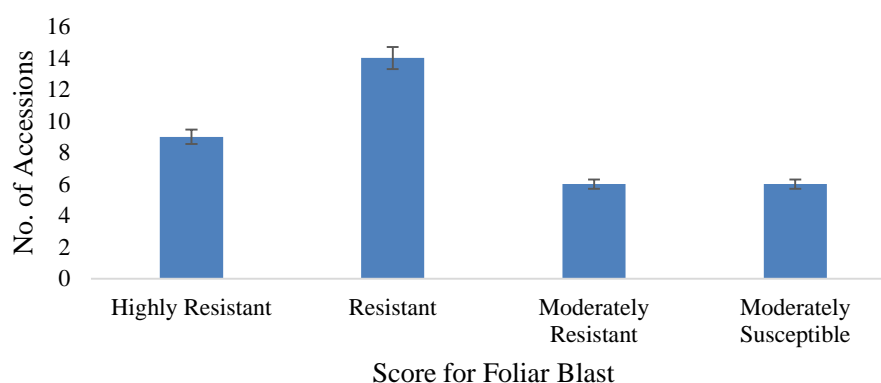
Figure 4. The score of foliar blast disease among the selected rice germplasms which are available in the Genesys Catalog.
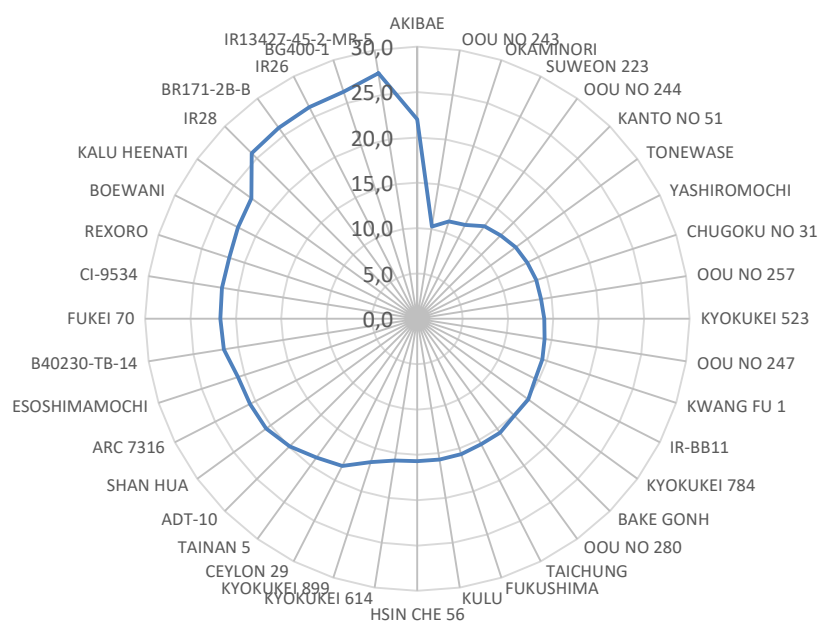
Figure 5. Data on amylose content which are available in the Genesys Catalog.

4. DOI implementation on rice germplasm in MARDI

MARDI was invited to be involved in the Project "W3B-PR-29-Indonesia: Multicountry Construction of a Test Platform for the Development and Allocation of Unique Identifiers to Rice Germplasm, linking the MLS information infrastructure and the DivSeed repository". A total of 708 accessions of Malaysian rice germplasms had been registered for DOIs (Figure 6). The accessions include 45 accessions of MARDI varieties, 591 accessions of Malaysian landraces and 72 accessions of introduced varieties from International Rice Research Institute (IRRI).

The released varieties consisted of the first variety released in 1964 until 2017. In Malaysia, the breeding program was started by the Department of Agriculture (DOA) in year 1960s. The first variety named Malinja was released in 1964 for double cropping planting season [5]. Malinja was derived from the cross between Siam 29 (landraces with photoperiod sensitivity) and Pebifun (introduced variety with non-photoperiod sensitivity). The variety has maturity days about 150 days from sowing. Mahsuri was the second released variety which was also popular during that period, but it had two serious weaknesses, i.e. susceptible to blast disease and lodging problem. Improvement was made to the variety through hybridization. Remadja and Sigadis, two varieties from Indonesia, were used as blast resistant donor varieties. Meanwhile, a few introduced varieties such as IR8, IR119 and IR2070 were used for the improvement of plant type and lodging problem in Mahsuri [5]. The last variety released in 1970s was Bahagia. The strain was selected from segregating IF5 (F4) population brought from IRRI. After four years of selection, the variety was released in 1968 [5]. Bahagia had good physical grain quality, medium height and relatively good yield.

The varietal improvement program in Malaysia has begun from the introduction of exotic elite breeding lines and varieties started with IR8. IR8 was introduced in 1966 for tropical irrigated lowlands. It was the first semi-dwarf and high yielding rice variety [6,7]. The variety increased yield potential of irrigated rice from 6 to 10 ton per ha in the tropics. In MARDI, the variety was used extensively in crosses either to improve the variety itself or to increase the yielding potential of other varieties [5]. Between 1968 and 1969, IR8 had been used in 53 crosses either with landraces, released varieties, varieties with resistance to pests and diseases or varieties with good grain quality. From

1972 to 1978, all the MARDI released varieties derived from these crosses namely were Murni, Masria, Jaya, Sri Malaysia I, Sri Malaysia II and Pulut Malaysia I.

Figure 6. List of Malaysia rice germplasm registered for DOI in Global Information System (GLIS).

Starting in 1978, the breeding program established new objectives based on current problems occurred in rice plantation. The objectives were targeted to solve the problems on pests and diseases, problem soils, drought and semi-deep water, rainfed and regional specificity. The introduced varieties used to achieve these objectives were from India, Bangladesh, Thailand, Indonesia, Philippines (including IRRI), Japan, Korea, China and others. Many landraces were also used as donors, especially for blast resistance, rice varieties for rainfed areas, and to broaden the genetic base. The local semidwarf mutants derived from mutation technique were also used as donors of dwarf genes.

Till date, MARDI has released three white glutinous rice varieties, one black glutinous rice variety, two Clearfield rice varieties that carry the resistant gene to imidazolinone to prevent weedy rice problems, four quality rice varieties, one aerobic rice variety and the remaining were ordinary rice. The ordinary rice usually consists of high yielding traits such as many tillers, long panicle, heavy grain weight and also resistant to pests and diseases.

Figure 7. The upland and rainfed rice cultivation in Malaysia.

Besides, selected valuable rice germplasm conserved in MARDI Rice Genebank were also successfully registered for DOIs, especially the Malaysian landraces. The landraces were collected throughout the country and is maintained and conserved in seeds genebank [1]. The landraces are important as a source of new genes [1] and essential for the development of superior recombinants in crop improvement and development of new rice variety [1,8]. The landraces are divided into two groups based on planting practices namely upland (Figure 1) and rainfed lowland (Figure 2).

Upland rice can be found in some area in Peninsular Malaysia but commonly found in Sabah and Sarawak. It is usually cultivated for home consumption by rural people in the states [9]. Landraces are mostly photoperiod sensitive, lengthy maturation, tall plant types, susceptible to lodging and less responsive to fertilizer. In MARDI, landraces have been successfully used in pre-breeding to generate desirable genotypes for new rice varieties. There are several outstanding landraces used in rice improvements such as Radin Goi, Pongsu Seribu 2, Mayang Ebos, Siam 29, Pulut Sutera, Tangkai Rotan, Engkatek and Secupak. All the registered accessions for DOI are stated under Annex 1 crop. They can be accessed by the contracting parties of the Treaty for research and crop improvement to achieve global food security and promoting sustainable agriculture in each country.

5. Challenges handling the toolkit DB for DOI registration
The DOI Toolkit DB was developed using Linux and the system ran using Java application. The setup of the Toolkit required large storage capacity of the computer system up to 8,000 MB or 4 GB. Thus, only with the large storage capacity of the system may ensure the success of the installation process. The Toolkit also requires the users to be familiarized with the SQL for running the trigger for registration. The trainees during the workshop were mostly the genetic resource persons who faced problems in handling the SQL for running the triggers. Further improvement of the system with the most effortless installation process and friendly use of DOI Toolkit for registration may be required for future and global use.

6. Conclusions
The involvement and contribution of MARDI in both projects showed the commitment of MARDI for data sharing mechanism established by the Treaty. The contribution will enhance rice breeding program and sharing valuable information on rice genetic resources with other countries for future research as well as for rice varietal development to achieve global food security and for promoting sustainable agriculture locally and globally.

7. References
[1] Shukor A R, Mohd Yusoff A, Norowi H and Shukor N 2009 Contribution of biodiversity towards sustainable agriculture development *Biodiversity and National Development Achivments, Opportunities and Challenges* p 20–31

[2] Anonim 1997 *Assessment of Biological Diversity in Malaysia* (Kuala Lumpur: Ministry of Science Technology and Environment)

[3] Tosiah S, Muhamad-Radzali M, Mohd-Shukor N and Salma I 2008 Information system for agrobiodiversity collections and conservation *Conference on Biodiversity and National Development* 28 May 2008 (Kuala Lumpur: The Legend Hotel)

[4] Lin M S 1991 Genetic base of *japonica* rice varieties released in Taiwan *Bot. Stud.* **56** 43–6

[5] Othman O, Alias I and Hadzim K 1986 Rice varietal development in Peninsular Malaysia (Kuala Lumpur)

[6] Khush G, Coffman W and Beachell H 2001The history of rice breeding: IRRI's contribution ed Rockwood W *Rice Research and Production in the 21st Century: Symposium Honoring Robert F Chandler Jr.* (Los Baños: International Rice Research Institute) p 117–35

[7] Anisuzzaman M, Kader M, Ali M, Haque M and Halder T 2016 Development of high yielding rice varieties for favorable ecosystem with 40% higher yield than the present variety: a review paper *Middle-East J Sci Res.* **24** 3644–53

[8]   Sohrabi M, Rafii M Y, Hanafi M M, Akmar S N and Latif M A 2012 Genetic diversity of upland rice germplasm in Malaysia based on quantitative traits *Sci. World J.* **2012** 1–9

[9]   Hanafi M M, Hartinie A, Shukor J and Mahmud T M M 2009 Upland rice varieties in Malaysia: agronomic and soil physico-chemical characteristics *Pertanika J. Trop. Agric. Sci.* **32** 225–46