NGS-BASED GENOMIC CHARACTERIZATION OF AGRICULTURALLY IMPORTANT PGR TO SUPPORT NATIONAL BREEDING PROGRAM

I Made Tasma

INTRODUCTION

Indonesia is recognized as the second richest megabiodiversity of the world. This includes plant genetic resources (PGR) of agriculturally importance. The PGR diversity determines the future of Indonesian agriculture. Few examples of PGR of Indonesian origins include rice, durian, banana, sugar cane, mangosteen and sweet potato. The PGR diversity needs to be explored to obtain genes, QTLs of interests for crop and animal breeding purposes. This can be done by the use of high throughput next generation sequencing (NGS) and high throughput SNP array technologies. Genes (DNA markers) are used for crop improvement in breeding programs. This paper describes the recent progress and achievement of plant genome re-sequencing projects at the Indonesian Agency for Agricultural Research and Development (IAARD). This manuscript describes the characterization of national agriculturally important crop species at genomic level based on whole genome sequencing technology using NGS platform. The crops characterized include soybean, maize, oil palm, cacao, chili pepper, potato, banana and *Jatropha curcas*.

A Brief Review of the IAARD Advanced Genomic Research Program

IAARD missions include developing superior crop cultivars using the available breeding techniques. The methods should be able to manipulate PGR collection richness in efficient and effective manners. This manipulation would be facilitated by the current genomic technology by using the high throughput genetic platforms. Since the end of 2010, the IAARD has been equipped with such high throughput sequencing platform (i.e. the NGS system, Illumina HiSeq2000) and also a high throughput SNP array reader (Illumina iScan). The objectives of genomic research of IAARD were to: (1) sequence and re-sequence national priority crops and animals; the sequence data will be used for gene and SNP discovery and SNP chip development; (2) dissect economically important genes (QTLs) of economically important traits (e.g. yield, biotic and abiotic stress tolerance, nutritional values, etc.) of the national priority crops and animals using high throughput marker e.g. SNP markers; and (3) develop superior lines (cultivars, superior animals) of priority traits (e.g. yield, meat production and quality) of the national priority crops and animals through molecular breeding techniques.

Table 1. Genomic data status of targeted crops and animals studied at the ICABIOGRAD (Tasma et al. 2012). The study was started in fiscal year of 2010.

Crop	Reference genome sequence	SNP marker	Research activities
Oil palm	Not available	Not available	Design marker
J. curcas	Not available	Not available	De novo sequencing
Cacao	Available	Not available	Design marker
Soybean	Available	Not available	Design marker
Rice	Available	Available	Association analysis
Cow	Available	Available	Association analysis
Maize	Available	Available	Association analysis
Banana	Available	Not available	Design marker
Potato	Available	Not available	Design marker
Chili pepper	Available	Not available	Design marker

The genetic platforms have been utilized to sequence, re-sequence, and genotype national high priority crops and animals (Table 1). The crops and animals studied included rice, soybean, cacao, oil palm, physic nut, banana, maize, potato, chili pepper and cows. For crops with no reference genome sequence yet available, a *de novo* sequencing and genome assembly project were performed. This includes *de novo* sequencing project of oil palm (as in the year of 2010 the genome reference sequence of oil palm was not yet available at that time) and *Jatropha curcas*. If genome reference sequences were available but SNP markers were not available, a re-sequencing project was commenced to obtain SNP databases of the crops. This was done for crops and animal such as soybean, cacao, banana, potato and chili pepper. For crops and animals with reference genome sequence and SNP markers available, a high throughput genotyping and trait association studies were accomplished. This mainly was done in rice, maize, and cows. The genomic status of the crop and animals studied at IAARD is as shown in Table 1.

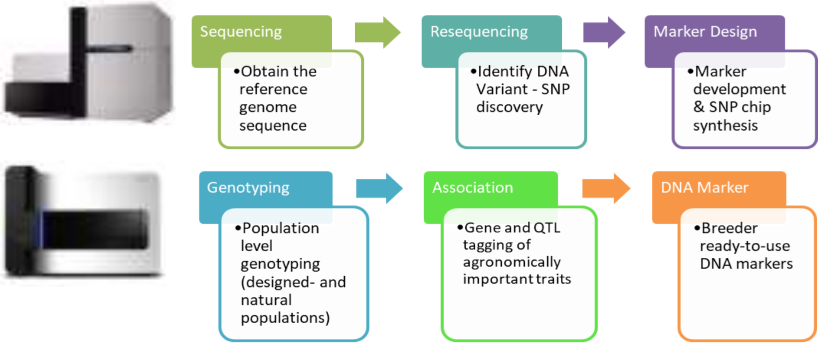


Figure 1. Strategy of advanced genomic studies at IAARD (Tasma et al. 2012; Tasma 2015; Tasma et al. 2018). Two high throughput genetic platforms (HiSeq2000 and iScan) have been applied in the research to support genomic-based breeding programs of priority crops at IAARD.

To meet the IAARD genomic research goals, we have designed a general strategic scheme of advanced genome research at ICABIOGRAD as shown in Figure 1. The NGS sequencing platform has been oriented to conduct *de novo* sequencing, re-sequencing, and genotyping by sequencing (GBS). While the high throughput SNP array reader is subjected for high throughput genotyping to expedite gene and QTL discoveries of important traits to be applied in a breeding program of the respective crops.

De Novo Sequencing of PGR Species to Develop Reference Genome Sequences

De novo sequencing is to define the genome sequence of a particular crop species the first time (Van et al. 2012; Danaval et al. 2012; Tasma 2017). *De novo* sequencing is intended to develop reference genome sequence of a particular crop species of interest. The reference genome will be used as a guidance to analyze genomic variations among individual genotypes of the species member using a simpler and cheaper method known as resequencing genome project (Tasma 2015; Tasma 2016).

A reference genome is required for each target species for genomics scientists to leverage the power of modern genome technologies in agricultural research. To produce and annotate the first genome sequence of any individual in a species, chromosomal DNA is broken up into billions of pieces, from 500 to 30.000 base pairs in length, and a complex set of specialized DNA libraries over many months are produced (Varshney et al. 2009; Wetterstrand 2011). High cost long read sequence data is then generated from the complex DNA libraries and a map of long contiguous chromosomal regions is reconstructed with a *de novo* sequence assembly computer. Long read sequences enable the computer programs to assemble through the tens of thousands of repetitive elements typically present in plant genomes, which in turn results in longer sequence scaffolds. Long sequence scaffolds have more utility to researchers than short sequence scaffolds. Once lengthy sequence scaffolds have been produced, expressed regions of the genome are typically annotated with millions of sequences generated from RNA libraries also using long read sequencing technologies. Similar to *de novo* genome sequencing, long read sequencing technology is also preferred over Illumina short read sequencing for the construction of the first-draft of the transcriptome, because long reads when assembled are more likely to result in full-length transcripts, which are most useful for gene model construction.

High quality (i.e., long and accurate) reference sequences are far more expensive to produce than genome re-sequencing data sets (Wetterstrand 2011). Moreover, research teams generally never complete reference projects but are engaged in a process of continually improving and releasing more consolidated and more accurate genome builds over time. For example, in the United States, even though the first reference genome sequence of man was completed in 2002, today an estimated 300 scientists are engaged in editing the reference human sequence, which is now in its 19th build. The more accurate the reference sequence is for a given species, the more powerful each low cost re-sequencing study becomes.

In researching a particular crop species, scientists often are required to produce 3 to 5 reference genome sequences from multiple species which are related to the crop (Varshney et al. 2009). For example, in maize (*Zea mays*), a reference genome sequence was produced from B73, MO17, as well as from the popcorn variety (Schnable et al. 2009). Likewise in rice (*Oryza sativa*), reference genome sequences have been produced from the niponbare, indica (Yu et al. 2002) and japonica (International Rice Genome Project) varieties. In cacao (*Theobroma cacao),* reference genome sequences have been produced from the Belizean Criollo (Argout et al. 2011) and Matina 1-6 cultivars. In Oil Palm, private groups have produced reference genome sequences from three *E. guineensis* palms with pisifera, dura, and tenera fruit types and from the related *E. Oleifera* oil palm. Reference genome sequences of several agriculturally important plant species have been reported (Table 2).

The IAARD have conducted *de novo* sequencing of two crop species, oil palm and physics nut (*Jatropha curcas*) to provide reference genome maps of the two crops (Tasma et al. 2014b). The oil palm genome reference map, however, was published in 2013 (Singh et al. 2013) and the analyzed sequence data obtained from three oil palm genotypes conducted by IAARD were used as re-sequenced data to discover genomic variations of the oil palm genomes. SNP and INDEL markers have been developed from the sequence data to support breeding programs of oil palm. *De* *novo* sequencing project of *Jatropha curcas* genome was conducted in collaboration with Seoul National University (Korea) and Kasersart University (Thailand). *J. curcas* genome reference map has been completed that becomes a high value resource to support breeding programs of this oil-bearing crop species. Millions of SNPs and Indels have been discovered by aligning the reference sequences with other genotypes of the Indonesian physic nut accessions. These would be very important marker resources for expediting physic nut breeding programs.

Table 2. Plant species of agriculturally importance in which their reference genome sequences have been available (Van et al. 2012; Danaval et al. 2012; Tasma 2015; Tasma et al. 2018).

Plant species	Scientific name	Genome size (Mbp*)
Rice	*Oryza sativa* ssp *indica*	430
	O. sativa ssp *japonica*	430
Soybean	*Glycine max*	1.115
Maize	*Zea mays*	2.300
	Zea mays spp. *parviglumis*	2.100
Sorghum	*Sorghum bicolor*	730
Cassava	*Manihot esculenta*	760
Papaya	*Carica papaya*	372
Cucumber	*Cucumis sativus*	367
Potato	*Solanum tuberosum*	844
Banana	*Musa accuminata*	472.2
Apple	*Malus domestica*	742.3
Strawberry	*Fragaria vesca*	240
Grape	*Vitis vinifera*	487
Watermelon	*Citrullus lanatus*	450
Melon	*Cucumis melo* L.	430
Tomato	*Solanum lycopersicum*	900
Green beans	*Phaseolus vulgaris*	486.9
Cacao	*Theobroma cacao*	430
Cotton	*Gossypium raimonddi*	750
Oil palm	*Elaeis guineensis*	1.800
Castor nut	*Ricinus communis*	350

**1 Mbp (Mega base pair) = 10^6 base pairs.*

Resequencing of PGR Genomes of Agriculturally Important Crops to Identify Genomic Variation

Once a set of reference genome exists, scientists can re-sequence a second or third individual from the same species using very low-cost short read technologies, such as the Illumina HiSeq-2000 platform (Varshney et al. 2009; Tasma 2016). Individual reads from a short read project are not necessarily assembled to one another, but rather are mapped back to the reference genome to identify genetic differences that exist between the reference sequence and the sequence of the second individual. In this way a list of differences can be recorded, and these genetic differences are used to explain differences between varieties within a species, such as yield, disease resistance, flowering time, drought and flood tolerance etc. Re-sequencing projects can be used to survey variation with a species, to conduct genotyping by whole genome sequencing studies, to build high resolution genetic maps, and to generate genome wide association maps.

Genome re-sequencing analysis was done to identify DNA variations of national priority crops and animals, marker SNP collection of which, has not been available. Whole genome of various superior accessions of the priority crop species were re-sequenced and genome variations (e.g. SNP and INDEL) were identified by aligning the genome reference sequence with those of the re-sequence data of the crop species under study. The type of crop accessions sequenced was Indonesian accessions and varieties to discovered genes and QTL adaptable to Indonesian tropical climates. The collected SNPs and INDELs from the studies have been validated, maintained in a database and be used in marker discovery analyses. The validated SNP markers will be selected based on genetic merits and GR characterization and future breeding objectives. The selected SNPs then are arranged in condensed manner in a chip containing low-to-high density SNP markers. The markers selected in such a way to meet scientific needs for discovering genes and GR evaluation of the respective crop or animal under study.

IAARD has focused on nine priority crop species important for Indonesian agriculture. The crops under study included rice, maize, soybean, oil palm, cacao, banana, chili pepper, potato and physics nut (*Jatropha curcas*). The genetic materials used in the genomic analyses were the parental breeding materials (i.e. the parents of breeding programs that were used in cultivar development of the targeted traits in each crop or animal species). The genetic materials analyzed also included other genetic materials of specific interest for the respective RIs important for their future breeding programs. The RIs also develop mapping populations based on traits of interest and as the targeted trait improvement for each crop and animal species.

The number of genetic materials used in this study ranged from three to 20 genotypes of each respective crop and animal species. The genomic DNA of the selected genotypes was deeply-resequenced using an NGS HiSeq platform to result in high quality reads. Genome coverage of this whole genome sequencing study ranged.

from 27 to 1.185 genomes (Table 3). The high quality DNA sequences resulted from this study were aligned with the reference sequence of each respective crop and animal species. The alignment identified genome-wide DNA variations in each crop and animal species unde study. The alignment was done by using software Bowtie2 (Langmead and Salzberg 2012). The sequencing alignment was then followed by genome variation characterization by using computer software Mpileup within Samtools (Li et al. 2009). Anotation of the location and effect prediction of the identified SNP/INDEL were conducted using software snpEff (Cingolani et al. 2012). Finally, a genetic diversity analysis of the five genome sequences was conducted by using software DarWin (Perrier and Jacquemoud-Collet 2006). The types of genomic variations identified from this study included SNPs, INDELS, and SSRs. DNA variations discovered from this experiment were as presented in Table 3.

Table 3. DNA variations observed based on alignment results of the NGS-derived re-sequence data to the respective reference map of the crop and animal species under study (Tasma et al. 2014; Satyawan et al. 2014; Tasma et al. 2015; Tasma 2016; Tasma et al. 2018).

Crop/animal species	Genom size (Mb)*	Number of genotypes sequenced	Average genome coverage**	DNA variation frequency observed***	Total DNA variations observed		Total
					SNP	INDEL	
Soybean	1,115	5	33	288	2,690,000	459,000	3,150,000
Maize	2,300	4	27	634	2,805,145	240,526	3,247,037
Banana	472.9	14	47	57	5,159,450	571,885	5,731,335
Cacao	430	5	77	121	2,326,088	362,081	2,688,169
Oil palm	1,800	3	109	197	3,032,200	303,109	3,335,331
Chili pepper	2,649	6	69	101	24,760,787	1,316,987	26,077,774
Potato	844	6	27	160	4,166,472	339,372	4,505,844
Physic nut	450	3	125	1,185	92,007	nd	92,007
Total and range values		46	27-125	57-1,185	45,032,149	3,592,960	48,625,109

Gene-Based Genomic Variations Observed Among the Crop Species Under Study

Breeders are of utmost interest in obtaining DNA variations located within the exons (protein coding regions) to design functional markers useful for breeding programs. Two types of DNA variations were observed from exons. The frist was synonymous SNPs and INDELS (DNA variations that do not change proteins sequences due to the changes of bases in the site of observed SNP or INDEL). The second was non-synonymous SNPs and INDELS (DNA variations that change the amino acid sequences of the proteins due to the changes in the sites of the observed SNPs/INDELS).

The majority of the variations observed were located outside the genes, i.e. upstream, downstream, intergenic regions of the genes (Table 3). Among the DNA variations observed only a small portion was located within the exons (protein coding regions) (Table 4).

Table 4. Location of DNA variations observed within the genome of each crop species under study (Tasma et al. 2014; Satyawan et al. 2014; Tasma et al. 2015; Tasma et al. 2016).

Crop species	Number and percentage of genomic variation observed at different parts of the genome					
	Upstream gene	Downstream gene	Intergenic region	Intron	Exon	Other sites*
Soybean	55.870	1.174.917	1.393.216	381.701	95.154	53.343
(*Glycine max)*	(29.757%)	(26.367%)	(31.586%)	(8.654%)	(2.157%)	(1.209%)
Maize	1.698.426	1.772.088	2.563.804	663.812	230.412	349.914
(*Zea mays)*	(23.34%)	(24.35%)	(35.23%)	(9.12%)	(3.17%)	(4.81%)
Banana	869.078	855.128	774.327	391.795	48.382	34.535
(*Musa accuminata*)	(29.83%)	(28.761%)	(26.043%)	(13.177%)	(1.627%)	(1.161%)
Cacao	469.623	360.497	404.919	28.995	1.3462	500.348
(*Theobroma cacao)*	(29.93%)	(22.98%)	(25.81%)	(18.48%)	(0.86%)	(2.11%)
Oil palm	584.063	551.772	2.782.751	492.316	55.870	4.804
(*Elaeis guineensis)*	(13.062%)	(12.34%)	(62.232%)	(11.01%)	(1.249%)	(0.0107%)
Chili pepper	1.589.861	1.486.165	25.071.638	532.184	180.234	388.18
(*Capsicum annum)*	(5.44%)	(5.081%)	(85.72%)	(1.82%	(0.16%)	(1.33%)
Potato	1.625.886	1.932.527	2.450.034	1.123.125	654.923	369.317
(*Solanum tuberosum*)	(19.94%)	(23.70%)	(30.04%)	(13.77%)	(8.03%)	(4.527%)
Physic nut	1.225	1.492	90.652	326	500	1.407
(*Jatropha curcas)*	(1.33%)	(1.51%)	(96.06%)	(0.35%)	(0.53%)	(1.597%)
Range of percentage	1.33-29.83	1.51-26.37	25.81-96.06	0.35-25.68	0.53-8.03	0.01-4.80

Other sites included splice site acceptor, splice site donor, UTR 3 primer, UTR 5 primer.

Genomic Data Base Development on PGRs of Agriculturally Important Crop Species

The IAARD has developed genome browsers covering genomic and phenotypic data of rice, soybean, maize, cacao, oil palm, physic nut, chili pepper, potato and cows. The genome browsers contain genome wide resequencing data of each respective crop and animal species, SNP and INDELS information (gene or non-gen SNPs and INDELS) across the genome including the sequence information where the SNPs and INDELS located, etc. The genome browsers are being incorporated into the IAARD open-to-public genome database (http://genom.litbang.pertanian.go.id). This database is a very important resource for crop and animal scientists (geneticists, plant and animal physiologists and breeders) interested in finding particular genes and genetic markers of particular traits of the respective crop and animal species covered by the IAARD genome database. The genome resources should provide the basic data for developing breeding resources to expidate national breeding programs of the national priority crop and animal species.

Figure 3. Genomic data base containing millions of genetic variations resulted from genomic characterization of various agriculturally important crop species of national priority assessed with NGS Hiseq2000 (Rijzaani et al. 2016; Tasma et al. 2018).

Breeding Programs Based on PGR Genomic Data

The DNA markers (i.e. SNPs, INDELS) linked to the traits of interest would be used in breeding programs through marker-assisted selection (MAS) and genomic selection methods. Such methods would be able to exploit our plant and animal GR richness for developing superior plant and animals in a more efficient and effective manners that will support national programs on food self-sufficiency and food security for human kinds.

In plants and animal breeding programs, it has been commonly practiced to tag traits with molecular markers. More classical genomic tools have been successfully used to tag many important agronomic traits with more classical DNA markers such as RFLP, AFLP and SSR markers. The markers linked to the traits then are used in a molecular breeding program to select individual plants or animals having the trait of interest. Such marker-assisted selection (MAS), marker-assisted backcrossing (MAB) become very common to be applied in recent years to expedite breeding programs. More recently with the ability to map plant genome with dense molecular markers (i.e. SNPs), genomic selection (GS) technique becomes more popular for plant and animal breeders.

The GS is defined as the simultaneous selection for many thousands of markers, covering the entire genome so that all genes are expected to be in linkage disequilibrium with at least some of the markers (Meuwissen et al. 2001). The GS is very compatible with the recent genotyping technologies using NGS and high throughput SNP genotyping methods in which the marker assay can cover thousands and even few millions SNP markers. This technology becomes more feasible with a drop in genotyping costs, has GS become feasible, attracting the attention of crop and perennial plant breeders (Bernardo and Yu 2007). Genome-wide and cost-efficient marker systems are needed to apply GS in plants. Furthermore, dedicated GS breeding populations with effective population sizes (N_e) of around 20–50 individuals must be adopted to increase the extent of LD and fit into currently achievable genotyping densities of a few hundred markers per Morgan. In genetically heterogeneous populations still encompass large amounts of genetic variation for sustained genetic gains, recent experimental results in outbred plants, indicate that GS has better predictive ability than the classical polygenic model (Lee et al. 2008). A GS scheme has been tested in several crop plants where predictive equations for multiple traits are developed on the basis of high density genotyping and precise phenotyping of several hundred individual plants of a discovery population (training set) involving N_e in the range of 15–50. Selection accuracy of the predictive models is then assessed in a validation population, targeting the application of early GS at the individual plant level in progeny trials. This technique, once accomplished will be very accurate in selecting individual plants having the trait of interest and can be done in a faster, more predictable and more precise manner.

Future Perspectives

Future breeding program demands interesting challenges to develop crop varieties with multiple traits. This is due to the significant effect of global warming affecting crop cultivations in the field. Breeders should develop high productivity crops with lower water availability, tolerant to abiotic and biotic stresses (e.g., drought; aluminum and iron toxicity, flood; more dynamic disease and insect pest development and patterns; more nutritious crop products; etc.). Breeders need to develop crop cultivars in a time frame basis to provide plant materials to address the product demand and environmental problems. Indonesia has high PGR diversity that should be characterized in efficient, effective, and comprehensive manners to obtain genes of interest for plant breeding purposes. The NGS technology expedites gene and marker discoveries for traits of interest. Genomic-based breeding program should result in more variable crop cultivars and plant products that can be resulted in a faster and precise manner. In addition, the advanced technology should also be able to use of PGR richness in comprehensive manner in plant breeding program. Indonesian government, therefore, needs to allocate more resources to be allocated in the genomic-based plant breeding program to assure the national food self-sufficiency program be successful.

Conclusion

IAARD has sequenced the whole genomes of 46 genotypes covering eight agriculturally important crop species. A total of 48.625.109 DNA variations consisted of SNP, insertion and deletion (INDEL) were obtained from the NGS-based sequence data. The genomic variations have been deposited in the national genomic data base that can be accessed for research purposes. PGR characterization at the genome level is useful mainly in developing breeding resources to be used in gene discovery and marker development useful to expedite plant breeding program of agriculturally important crop species. The genome database containing genetic variations derived from 39 PGR of nine crop species will be useful for scientific as well as product development derived from the crop species. A more comprehensive genomic characterization of PGR would be more useful in supporting future breeding program. With the development and the lower cost of sequencing technology the genome sequencing will be affordable by the more common laboratories of the world. The NGS-based PGR characterization technology, therefore, will become routine activities in the future. The Indonesian government needs to be aware on this type of technology to be able to use its PGR richness in more efficient and comprehensive manners for breeding and product development purposes to support the food, feed and bioenergy self-sufficiency programs to be successful.

Acknowledgement

The author thanks the IAARD and ICABIOGRAD for funding the National Research Agricultural Genome Project (2010-2014), and the IAARD Genome Research Team for their involvement in developing such well genomic data of national agricultural important PGR.

REFERENCES

Argout, X., Salse, J., Aury, J.M., Guiltinan, M.J., Droc, G., Gouzy, J., Allegre, M., Chaparro, C., Legavre, T., Maximova, S.N., et al. [36 authors] (2011) The genome of *Theobroma cacao*. *Nature Genetics,* 4(2):101-108.

Bernardo, R. & Yu, J.M. (2007) Prospects for genome wide selection for quantitative traits in maize. *Crop Science*, 47:1082-1090.

Cingolani, P., Platts, A. & Coon M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, 6:80-92.

Dhanapal, A.P. (2012) Genomics of crop plant genetic resources. *Advances in Bioscience and Biotechnology*, 3:378-385.

Langmead, B. & Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods,* 9:357-359.

Lee, S.H., van der Werf, J.H., Hayes, B.J., Goddard, M.E. & Visscher, P.M. (2008) Predicting unobserved phenotypes for complex traits from whole-genome SNP data. *PLoS Genet*, 4:e1000231.

Li, H., Handsaker, B. & Wysoker, A. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics,* 25:2078-2079.

Meuwissen, T.H., Hayes, B.J. & Goddard, M.E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics,* 157:1819-1829.

Perrier, X. & Jacquemoud-Collet, J. (2006) DARwin software. Available at: http://darwin.cirad.fr/.

Rijzaani, H., Lestari, P., Priyatno, T.P. & Tasma, I.M. (2016) Pusat genom komoditas pertanian Indonesia. *Warta Balitbangtan,* 38(4):2-6.

Satyawan, D., Rijzaani, H. & Tasma, I.M. (2014) Characterization of genomic variation in Indonesian soybean (*Glycine max*) varieties using next-generation sequencing. *Plant Genetic Resources,* 12: S109-S113. Doi:10.1017/S1479262114000380.

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C. [152 authors] (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science,* 326:1112-1115.

Singh, R., Eng-Ti, L.L., Ooi, L.C.L., Ong Abdullah, M., Ting, N.C., Nagappan, J., Nookiah, R., Amiruddin, M.D., Rosli, R., Manaf, M.A.A., Chan, K.L., Halim, V., Azizi, N., Lakey, N., Smith, S.W., Budiman, M.A., Hogan, M., Bacher, B., Brunt, A.V., Wang, C., Ordway, J.M., Sambanthamurthi, R. & Martienssen, R.A. (2013) The oil palm shell gene controls oil yield and encodes a homologue of seedstick. *Nature,* 500:340-344. Doi:10.1038/naturel2356.

Tasma, I.M. (2014) Single nucleotide polymorphism (SNP) sebagai marka DNA masa depan. *Warta Biogen,* 10(3): 7-10.

Tasma, I.M., Satyawan, D., Rijzaani, H., Utami, D.W., Lestari, P. & Rosdianti, I (2012) Pembentukan empat peta genetik sawit, jarak pagar, padi, dan kedelai, serta identifikasi marka SNP kakao dan sapi. Laporan Akhir Penelitian APBN 2012. BB Biogen, Badan Litbang Pertanian.

Tasma, I.M., Rijzaani, H., Satyawan, D., Lestari, P., Reflinur, Rosdianti, I., Mansyah, E., Kirana, R., Kusmana, Pabendon, M. & Rubiyo, et al. (2014b) Analisis genom dan sidik jari komoditas pertanian strategis. Laporan Akhir Penelitian 2014. BB Biogen Bogor.

Tasma, I.M., Rijzaani, H., Satyawan, D., Lestari, P., Utami, D.W., Rosdianti, I., Purba, R., Mansyah, E., Sutanto, A., Kirana, R., Kusmana, Anggraeni, A., Pabendon, M. & Rubiyo. (2015) Next-gen-based DNA marker development of several importance crop and animal species. Presented at SABRAO 13th Congress and International Conference 14-15 September 2015, Bogor, Indonesia.

Tasma, I.M. (2016) Resekuensing genom, metode baru karakterisasi variasi SDG tanaman secara komprehensif mendukung akselerasi pemuliaan tanaman. *Warta Biogen,* 12(1):2-6.

Tasma, I.M. (2017) Sekuen genom acuan (*reference genome sequence*), kunci sukses program pemuliaan tanaman berbasis data genom. *Warta Biogen,* 13(1): 9-12.

Tasma, I.M., Sustiprijatno & Mastur (2018) Sekuensing genom untuk karakterisasi sumber daya genetik tanaman. In: *Pemanfaatan SDG dan Bioteknologi untuk Mendukung Pertanian Berkelanjutan.* Bogor: IAARD Press. p 83-110.

Van, K., Rastogi, K., Kim, K.H. & Lee. S.H. (2013) Next-generation sequencing technology for crop improvement. *SABRAO Journal of Breeding and Genetics,* 45(1):84-99.

Varshney, R.K., Nayak, S.N., May, G.D. & Jackson. S.A. (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends Biotechnol,* 9:522-530.

Wetterstrand, K.A. (2011) DNA Sequencing costs: data from the NHGRI large-scale genome sequencing program. Available at: http://www.genome.gov/sequencingcosts, Accessed March 1, 2011.

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Yang, X., Cao, M., Liu, J., Sun, J., et al. [58 authors]. (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science, 296(5565):79-92. Doi:10.1126/science.1068037.

TOWARDS BUILDING THE BIG DATA OF SHALLOT: PHENOMIC, METABOLOMIC AND GENOMIC

Awang Maharijaya, Evriana Eka Pratiwi, Marlin, Lina Herlina, Dyah Kurnianingtyas, Ita Aprilia, Diny Dinarti, Reflinur, Suryo Wiyono, Agus Purwito, and Sobir

INTRODUCTION

Shallot (*Allium cepa* var *ascalonicum*) is an important commodity, especially in Asia including Indonesia. Shallot is consumed in both fresh and processed forms. The use of shallot has been widely known as a flavoring in food or used as an ingredient in herbal medicines (Grubben and Denton 2004). Demands for shallots in Indonesia and in the international markets continue to increase from year to year. Therefore efforts are needed to fulfill these needs through increased production. Unfortunately, in Indonesian the production of shallot is heavily reliant on very limited number of varietieslike Bima Brebes. In fact, the yield of this variety has been decreasing from year to year. In addition, the use of seed bulb for propagation has contributed to the narrowing genetic diversity of shallot in Indonesia.

Breeding is a very urgent thing to do and genetic diversity studies are important in optimizing plant-breeding programs toward high yielding and high bulb quality shallot varieties. Genetic diversity could be analyzed using morphological characters, chemical characters, and molecular markers (Wahyuni *et al*. 2013). The main objective of this study was to get the information regarding the genetic diversity of 59 shallot genotypes and initiating the construction of big data for shallot in Indonesia. Thirty five morphological traits (macroscopic and microscopic), ability to produce flower, untargeted metabolomics approach using CG-MS and LC-MS, and some random and gene analog-based markers were used to portray the diversity of our shallot collection. Further, the acquired data were used to identify flowering related genes in shallot and to identify the defense mechanism in shallot against pathogens.

All Genetic Markers Reveal the High Genetic Diversity of Shallot in Indonesia

Our results showed that our shallot genotypes have high genetic diversity (Figure 1, Figure 2, and Figure 3). Cluster analysis of 40 genotypes of shallot divided them into two main groups based on the coefficient of dissimilarity of 0.59 (21 morphological traits) and also two main groups based on the coefficient of dissimilarity of 0.48 (14 morphological bulb traits). Principal component analysis based on morphological characters showed that there were three main components that could explain 62.10% of the total diversity.

Untargeted metabolomic analysis of shallot bulb using GC-MS can detect a total of 326 metabolite compounds. Based on the heatmap of 326 metabolites, the metabolites are grouped into two groups, group A and group B (Figure 3). Group A consists of three metabolite compounds, namely *cycloartenol* (*triterpenoid*), *palmitic acid* (*fatty* acid), *octadecadienoic acid* (*fattyacids*). Group B comprises compounds that are only found in one or several genotypes. The heatmap divided shallot genotypes into two groups. Group 1 contains genotypes which only have *palmitic acid* and *octadecadienoic acid* compounds. While group 2 contains genotypes which only have *palimitic acid* and *cycloartenol*. Some sulfur compounds found in some genotypes in this study were *methylsulfanyl-4,5,6,7-tetrahydro-benzo[c]thiophene* (BM26), *nonadecane* (BM29), *allyl-dimethylcyclopropane* (BM45), *lanosterol* (BM47), *stearic acid* (BM60) and *propyl alcohol* (BM21). Cluster analysis classified the genotypes into two main groups with a dissimilarity coefficient of 0.41. Eigenvector of the four main components could be used to reduce the 229 original characters to 24 main characters.

We also found that not all of shallot genotypes can naturally produce flower. Therefore more efforts are needed to induce flowering ability in those genotypes. Flowering is important to transfer desirable traits from one genotype to others during the breeding process. Moreover, the ability to produce flower and TSS (True Seed of Shallot) is also esential during the germplasm collection and preservation. So far, we could only store 18% shallot genotypes as TSS (Table 1), while the rest were stored as bulbs. The bulbs always cause problem since they can only be stored in relatively short periods.

Resistance gene analog-based markers could also be used to support the morphological and metabolites characterization in evaluating the genetic diversity of shallot (Herlina et al. 2018). The diversity and population structure of shallots has also been detected by resistance gene-derived markers. Six polymorphic *R* gene-derived markers (Acepa-1, Acepa-2, Acepa-3, Acepa-4, Acepa-5 and Acepa-6) were developed and applied to analyse the genetic diversity of shallots in Indonesia. These loci produced a total of 963 alleles with an average of 0.51 ± 0.77 alleles per marker. The polymorphic information content value ranged from 0.2784 to 0.5236 with an average of 0.36015 (Herlina et al. 2019).

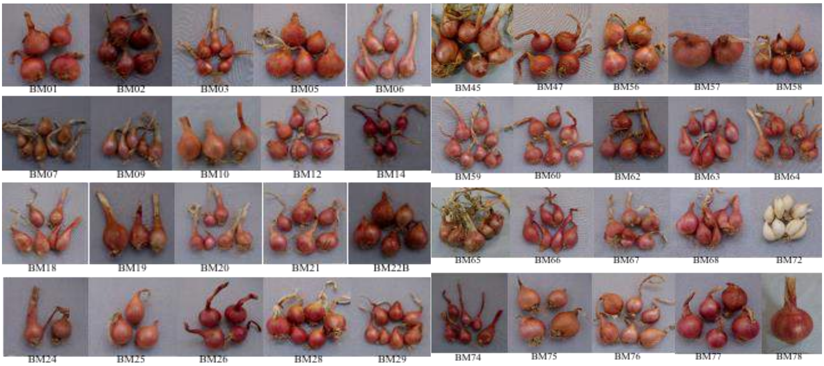


Figure 1. Morphological diversity of shallot bulb in our collection.

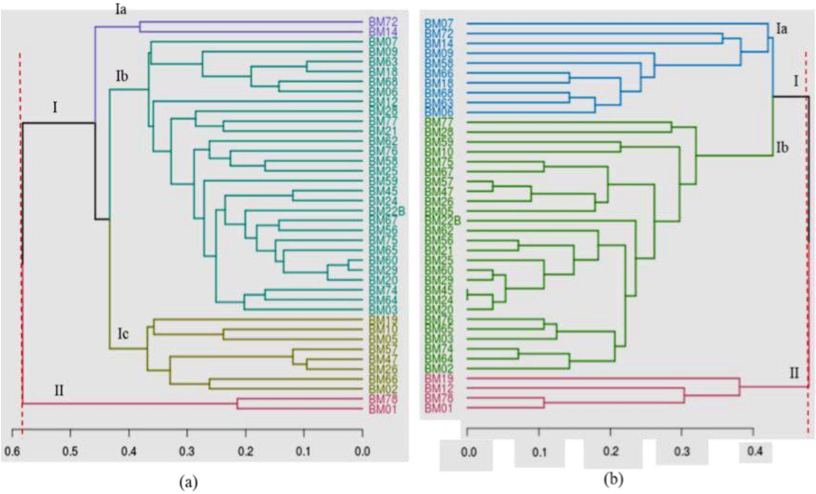


Figure 2. Cluster analysis of shallot diversity based on morphological and molecular characters.

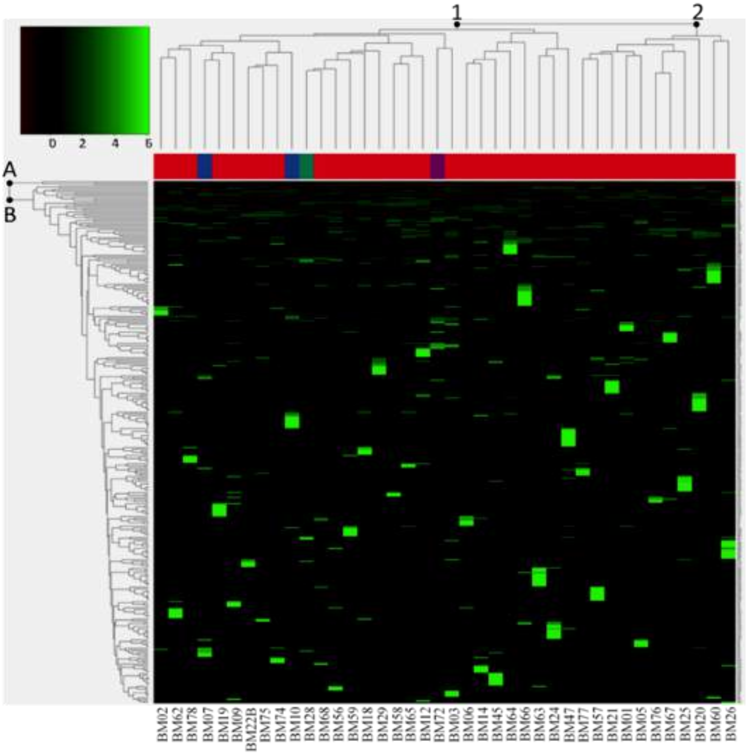


Figure 3. Cluster analysis of shallot diversity based on metabolite contents.

Table 1. Storage systems of shallot genotypes in our collection.

Geno	Collection material		Geno	Collection material		Geno	Collection material	
	Bulb	TSS		Bulb	TSS		Bulb	TSS
BM 01	√	-	BM 35	√	√	BM 60	√	-
BM 02	√	√	BM 36	√	-	BM 61	√	-
BM 03	√	√	BM 40	√	-	BM 62	√	-
BM 05	√	√	BM 41	√	√	BM 63	√	-
BM 06	√	-	BM 42	√	-	BM 64	√	-
BM 07	√	-	BM 43	√	-	BM 65	√	-
BM 08	√	-	BM 44	√	√	BM 66	√	-
BM 09	√	-	BM 45	√	-	BM 67	√	-
BM 10	√	-	BM 46	√	-	BM 68	√	-

Table 2. Continue.

Geno	Collection material		Geno	Collection material		Geno	Collection material	
	Bulb	TSS		Bulb	TSS		Bulb	TSS
BM 12	√	√	BM 47	√	-	BM 69	√	-
BM 14	√	-	BM 49	√	-	BM 70	√	-
BM 15	√	√	BM 50	√	-	BM 71	√	-
BM 16	√	-	BM 51	√	-	BM 72	√	-
BM 18	√	-	BM 52	√	-	BM 73	√	-
BM 19	√	√	BM 53	√	-	BM 74	√	-
BM 20	√	-	BM 54	√	-	BM 75	√	-
BM 21	√	√	BM 55	√	-	BM 76	√	-
BM 24	√	-	BM 56	√	-	BM 77	√	-
BM 25	√	√	BM 57	√	-	BM 78	√	-
BM 26	√	-	BM 58	√	-	BM 79	√	-
BM 29	√	-	BM 59	√	-			

Molecular Diversity of the Flowering Related Gene (Leafy) on Shallot

Since the flowering ability of shallot is very important for breeding purpose and propagation, we turned our focus to this aspect. Marlin et al. (2018) used the previously described material to study flowering in shallots. Flowering initiation in shallot has predictably been associated to the function of LEAFY (LFY) gene. This gene has been reported as the flowering meristem identity, and a potential indicator of plant flowering ability. Marlin et al. (2018) identified the diversity of shLFY (shallot-LFY) gene in five shallot genotypes in relation to their flowering pattern, and compared them to homologous sequences from Allium relatives. Genomic DNA from the 5 genotypes was amplified using novel primers (F/5′-GACCCAACAGACCCCTAC-3′ and R/5′-TTAGAACATAGAGGACACAGAG-3′), resulting in shLFY gene sequences with fragment lengths between 1.161 and 1.253 bp. These fragments were found to contain 3 exons with 2 introns.

The sequences of shLFY from 5 shallot genotypes confirm the existence of genetic polymorphism among shallot genotypes. The result of sequence analysis based on genomic and amino acid sequences identified 3 groups of shLFY genes in shallot. The first group consists of a single member, bm1LFY (Bentanis), which corresponds to naturally flowering type. The second group contains bm2LFY (Bima Brebes) and bm4LFY (Tajuk), which correspond to inducible flowering type. The third group comprises bm3LFY (Ilokos) and bm5LFY (Sumenep), corresponding to the non-flowering shallot type. The ShLFY sequence encodes a putative protein of 363 amino acids, with ~ 99% homology to the Allium cepa LEAFY and > 95% homology to LEAFY proteins from other higher plants. The shLFY protein in the 5 shallot genotypes also showed homology with FLORICAULA/LFY protein from referenced Allium relatives.

Identification of Defense Mechanisms Against Stress in Shallot

Using the previously described materials, we identified some important resistance traits in shallot, such as resistance to *Fusarium oxysporum*, *Colletotrichum gloeosporioides* and adaptability in tidal swampland. Aprilia et al. (2019) found that two shallot genotypes, Batu Ijo and Rubaru or Sumenep, were resistant to *F. oysporum.* Recently, Kurnianingtyas et al. (2019) identified that the Sumenep variety was the most resistant to *C. gloeosporioides,* with a disease severity of 30.19%, while the Biru lancor variety was the most susceptible with a disease severity of 95.05%. Laboratory tests indicated that the resistance to *C.gloeosporioides* might be related to the thickness of the palisade tissue (R=0.80). This research also showed a possible relationship between resistance to *C. gloeosporioides* and metabolite productions. We detected that carbamic acid, an elicitor compound in the jasmonic acid transduction pathway, was highly correlated with the resistance. Similar approach has also been done by Galingging et al. (2018) to detect metabolite contents of shallot genotypes that might be related to the adaptation of shallot to extreme conditions like tidal swampland.

Concluding Remarks and Future Perspective

Genetic diversity of shallot in Indonesia can be considered as high. This diversity should be exploited further in breeding programs to make shallot production more sustainable by reducing chemical usages for controlling pest and diseases in Integrated Pest Management system (Maharijaya and Vosman 2015). Our materials are available in the form of commercial and near commercial varieties of shallot. Thus, we can avoid breeding constraints introduced by the use of wild relatives for breeding, such as hybrid sterility and low crossing ability, as well as the retention of undesirable agronomic traits (Hajjar and Hodgkin 2007). However, there are still some important issues to be solved, such as induction of flowering ability from desirable genotypes for breeding purposes. Networking and collaboration are urgently needed and databases should be developed to be made available and accessible for related institutions.

Acknowledgment

This work was financially supported by grants from the Ministry of Research, Technology and Higher Education of the Republic of Indonesia.

REFERENCES

Aprilia, I., Sobir, S., Wiyono, S. & Maharijaya, A. (2019) Keragaman genetik dan ketahanan terhadap penyakit layu fusarium (*Fusarium oxysporum* f.sp cepae) bawang merah (*Allium cepa* L. var. aggregatum) Indonesia. *Jurnal Hortikultura Indonesia* (Under review).

Galingging, R.Y., Sobir, S., Aisyah, S.I. & Maharijaya, A. (2018) GC-MS profiling of volatile compounds from fifteen different varieties of Indonesian shallot grown in tidal swampland. *Rasayan J Chem*, 11(2):575-581.

Grubben, G. & Denton, O.A. (2004) *Plant resources of tropical Africa 2; Vegetables*. Wageningen, PROTA Foundation: ckhuys Publishers.

Hajjar, R. & Hodgkin, T. (2007) The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica*, 156(1-2):1-13.

Herlina, L., Reflinur, R., Nugroho, K., Terryana, R.T., Sobir, S., Maharijaya, A. & Wiyono, S. (2018) Genetic diversity analysis using resistance gene analog-based markers to support morphological characterization of shallots. *Jurnal AgroBiogen*, 14(2):65-74.

Herlina, L., Reflinur, R., Sobir, S., Maharijaya, A. & Wiyono, S. (2019) The genetic diversity and population structure of shallots (*Allium cepa* var. aggregatum) in Indonesia based on R gene-derived markers. *Biodiversitas Journal of Biological Diversity*, 20(3):696-703.

Kurnianingtyas, D., Sobir, S., Wiyono, S. & Maharijaya, A. (2019) Resistance to *Colletotrichum gloeosporioides* Penz in shallot (*Allium cepa* var ascalonicum). *Crop Protection* (Submitted).

Maharijaya, A. & Vosman, B. (2015) Managing the colorado potato beetle; the need for resistance breeding. *Euphytica,* 204(3):487-501.

Marlin, M., Maharijaya, A., Purwito, A. & Sobir, S. (2018) Molecular diversity of the flowering related gene (Leafy) on shallot (*Allium cepa* var. aggregatum) and allium relatives. *SABRAO Journal of Breeding & Genetics,* 50(3):313-328.

Wahyuni, Y., Ballester, A.R., Tikunov, Y., de Vos, R.C., Pelgrom, K.T., Maharijaya, A., Sudarmonowati, E., Bino, R.J. & Bovy, A.G. (2013) Metabolomics and molecular marker analysis to explore pepper (*Capsicum* sp.) biodiversity. *Metabolomics,* 9(1):130-144.

CHAPTER 4.
DIGITAL OBJECT IDENTIFIER FOR PLANT GENETIC RESOURCES MANAGEMENT

The Digital Object Identifier (DOI) is defined as a character string used to identify intellectual property in the digital environment. The DOI system was established to provide a framework for managing intellectual content, including activities such as linking users to content owners, facilitating electronic commerce and enabling automated copyright management for all types of media. A DOI name is permanently assigned to an object, to provide a persistent link to current information about that object, including where the object, or information about it, can be found. The DOI system consists of four components: identifier, resolution, metadata or description and policy. The value of DOI system lies in its combination of resolution, metadata and policy.

The Global Information System (GLIS) is established by the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) of the United Nations Food and Agriculture Organization (UN-FAO) to facilitate agricultural researchers, plant breeders and farmers to access relevant information related to plant genetic resources for food and agriculture (PGRFA) around the world. GLIS is expected to become a one-stop shop for easy access to information on seeds and other crops materials for research, training and plant breeding. On the other hand, the development and promotion of the use of the established system of DOI is also needed as several communities have highlighted the importance of creating and adopting Permanent Unique Identifiers to improve the identification of PGRFA. It is expected that DOI will be the international standard for identifying and documenting PGRFA materials uniquely and permanently, as well as facilitating data interoperability among different systems.

Multi-country Construction of a Test Platform for the Development and Allocation of Unique Identifiers for Rice Germplasm is a collaborative project between the Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD), the Indonesian Agency for Agricultural Research and Development (IAARD), ITPGRFA-FAO Rome and IRRI in the Philippines. It was financed by Benefit Sharing Fund of the ITPGRFA and was intended to promote the use of DOI for rice germplasm management and exchanges among participating countries.

DOI can be used as an identifier for PGRFA material managed by individuals, groups and organizations. DOI identifies PGRFA materials, not the associated data. Therefore, if the data associated with the PGRFA materials changes, the holder must correct the data without changing the DOI. DOI also has direct association with the holder of PGRFA materials. When a material is transferred, assigning a new DOI is usually recommended because the material is potentially placed under different legal and quality management conditions. However, when the recipient publishes results, they must refer to the recipient's DOI, not the provider's.

It is important to include basic data or information related to PGRFA materials as well, which is known as DOI metadata. DOI metadata is essential for identifying PGRFA materials and providing an overview of the nature or category of the PGRFA materials. Adopting the DOI does not mean replacing existing systems with a new one. Both DOI and local identifiers can still be used to manage the identity of PGRFA materials, as well as becoming main references in online publications and articles. However, if the PGRFA material is transferred across different organizations, then local identifiers will not be sufficient to be used as material identifiers. A unique global identifier system such as DOI is preferable in such cases to facilitate access globally. Once the PGRFA material is given a DOI, the DOI metadata of this material will be uploaded to GLIS and can be accessed by global users. This is one of the advantages, where the assignation the DOI will significantly improve the visibility of PGRFA-related data.

Although tested for rice germplasm as a pilot project, the DOI registration for other PGRFA is the same and straight forward. Several countries have registered non-rice germplasm from their gene bank successfully. The DOI system also improved the plant genetic resources management since it will also keep track of the movement of accessions in the genebanks. In addition, registration of the accession in the DOI system for some countries help them to consolidate their accession since each DOI number should serve as unique identifier for each accession in the collection.

ACCELERATING DOI REGISTRATION OF PGRFAs FROM ASIAN AND AFRICAN COUNTRIES THROUGH INTEGRATION-TOOLKIT DISSEMINATION

Hakim Kurniawan, Nurul Hidayatun, and M. Sabran

INTRODUCTION

The DOI is defined as a character string used to identify intellectual property in the digital environment (Information Standards Quarterly 2004). The DOI system was established to provide a framework for managing intellectual content, including activities such as linking users to content owners, facilitating electronic commerce and enabling automated copyright management for all types of media. A DOI name is permanently assigned to an object, to provide a persistent link to current information about that object, including where the object, or information about it, can be found (Paskin 2010). The DOI system consists of four components: identifier, resolution, metadata or description and policy. The value of DOI system lies in its combination of resolution, metadata and policy (Chandrakar 2006).

The Global Information System (GLIS) is established by ITPGRFA of the United Nations (FAO) to facilitate agricultural researchers, plant breeders and farmers to access relevant information related to plant genetic resources for food and agriculture (PGRFA) around the world. GLIS is expected to become a one-stop shop for easy access to information on seeds and other crops materials for research, training and plant breeding. On the other hand, the development and promotion of the use of the established system of Digital Object Identifiers (DOI) is also needed as several communities have highlighted the importance of creating and adopting Permanent Unique Identifiers for improved identification of PGRFA. It is expected that DOI will be the international standard for identifying and documenting PGRFA materials uniquely and permanently, as well as facilitating data interoperability among different systems.

Multi-country Construction of a Test Platform for the Development and Allocation of Unique Identifiers for Rice Germplasm is a collaborative project between the Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD), the Indonesian Agency for Agricultural Research and Development (IAARD) Indonesia, ITPGRFA-FAO Rome and IRRI in the Philippines. It is financed by Benefit Sharing Fund and intended to promote the use of DOI for rice germplasm management and exchanges among participating countries. To facilitate DOI batch-registration of large collections of PGRFA, an Integration-Toolkit, is developed and disseminated during the implementation of the project.

DOI for Easier Management of PGRFAs

DOI can be used as an identifier for PGRFA material managed by individuals, groups and organizations. DOI identifies PGRFA materials, not the associated data. Therefore, if the data associated with the PGRFA materials changes, the holder must correct the data without changing the DOI. DOI also has direct association with the holder of PGRFA materials. When a material is transferred, assigning a new DOI is usually recommended because the material is potentially placed under different legal and quality management conditions. However, when the recipient publishes results, they must refer to the recipient's DOI, not the provider's.

It is important to include basic data or information related to PGRFA materials as well, which is known as DOI metadata. DOI metadata is essential for identifying PGRFA materials and providing an overview of the nature or category of the PGRFA materials. Adopting the DOI does not mean replacing existing systems with a new one. Both DOI and local identifiers can still be used to manage the identity of PGRFA materials, as well as becoming main references in online publications and articles. However, if the PGRFA material is transferred across different organizations, then local identifiers will not be sufficient to be used as material identifiers. A unique global identifier system such as DOI is preferable in such cases to facilitate access globally. Once the PGRFA material is given a DOI, the DOI metadata of this material will be uploaded to GLIS and can be accessed by global users. This is one of the advantages, that assignation the DOI will significantly improve the visibility of PGRFA-related data.

Important Data for Assignation of DOI

The object associated with a DOI name is described unambiguously by DOI metadata. Assignment of a DOI name requires the registrant to record metadata describing the object to which the DOI name is being assigned (Paskin 2010). DOI metadata is essential for providing a basic description of the characteristic of the PGRFA materials, thus users can easily screen and select based on their own needs. There is a minimum requirement for data sets, called descriptors, which need to be provided for assignation of DOI.

Descriptors associated to the DOI in GLIS can be categorized as core descriptors and additional descriptors. Core descriptors are data sets which are essential for GLIS to perform its main functions. Parts of the core descriptors are fundamental for assignation of DOI (mandatory descriptors), while others should be provided if available (highly-recommended descriptors) to get the advanced functions of GLIS. On the other hand, additional descriptors are complementary data sets which can be relevant depending on some values inputed in the mandatory descriptors (Alercia et al. 2018).

DOI Registration of PGRFAs

In general, there are 3 methods to do DOI registration i.e. manually on the GLIS web site (suitable for small collections or editing), using Excel tables (for medium size collections), and using the XML protocol (for all types of collections). Manual registration through website supports registration and update, whereas Excel tables supports registration but not update.

Contrary to manual registration and Excel batch registration options, the XML protocol offers greater flexibility. It supports registration and update, offers a system-to-system integration option, contains simple message structure, easy to implement and maintain, reduces human errors and minimizes human intervention. This method also has an extensive data validation and reporting feature, with very fast running time during the execution (over 3 messages per second). Nevertheless, stakeholders are welcome to implement their own software layer, or adopt the Toolkit to minimize the effort.

Assignation of DOI to PGRFA typically involves the following steps (Kurniawan et al. 2019):

1. Installation of the Toolkit in a local machine and configure it. The latest version of Toolkit v2.0.3 offers much simpler procedures for installation and configuration. The configuration includes creation of a local database to be used for storing PGRFA data to be uploaded into GLIS and obtain DOI after the registration is successful.

2. Installation of supporting software. The Toolkit is a Java-based application; therefore, it needs Java Runtime Environment to run properly. The other essential software is database management software such as DBeaver or My SQL Workbench. That software is all opensource and can be freely downloaded from their official websites.

3. Data preparation in a standard datasheet in Excel format. All required data regarding PGRFA is compiled and has to be inputted according to the standard datasheet.

4. Data importing and mapping. Data importing is conducted to move the validated PGRFA data from Excel format into the local database managed by the Toolkit. Data mapping is conducted as it is necessary for the additional data to be moved into particular tables in the database.

5. Deploying the Toolkit and starting to upload the PGRFA data for DOI registration.
6. Store the registered DOI for PGRFA into local database.

Integration-Toolkit Development and Dissemination

The Integration Toolkit, henceforth abbreviated as Toolkit, was developed to help those unwilling or unable to implement their own XML solution and implement the XML message processing. The Toolkit uses JDBC database as data source and destination, giving it more flexibility to manage PGRFA datasets as well as obtained DOI. The Toolkit is also expected to become a bridge to guarantee compliance with GLIS XML Protocol and compatibility with any future GLIS updates such as transfer transaction and SMTA reporting.

The Toolkit comprised two main components: the database containing the information to be provided to GLIS and for storing the responses from GLIS and the XML communication layer that formats the information in the database into XML messages that GLIS can accept and respond to. It is important to note that the Toolkit performs no validation on the data that are read from the database, so mistakes like duplicate materials will be processed as well. If there is any error, it will be GLIS that will return the corresponding error messages.The initial version of Toolkit v1.0.0 was released in 2017. Testing of the Toolkit was carried out at IRRI, Los Banos, Philippines in October-November 2017. Two ICABIOGRAD staffs worked together with staffs from IRRI's Crop Research Informatics Laboratory. During the testing of the Toolkit, a total of 849 accessions of Indonesian rice were registered to obtain DOI.

After the Toolkit testing and implementation in IRRI, a newer version of Toolkit v1.0.1 was released. Toolkit v1.0.1 was disseminated at the 1st DOI International Training Course in Bogor, Indonesia, at 23 April – 1 May 2018. In this training course, 16 participants from 7 countries (Indonesia, Malaysia, Philippines, India, Bhutan and Zambia) were present. The training aimed to socialize the Toolkit v1.0.1 software for batch-registration of DOI on rice accessions. During the training, registration of DOI rice collections from each participating country was also conducted. It is expected that the DOI registration of rice collections will continue to be carried out by each participant after they returned to their respective institutions. Some participants still encountered problems during the implementation of Toolkit v1.0.1. There were two common problems faced by the participants during the implementation of DOI registration using Toolkit v1.0.1:

1. Toolkit v1.0.1 can only be run in Linux operating system and some of its supporting software demand very high storage space and computer memory capacity
2. Toolkit v1.0.1 is executed by using command prompts, which is quite difficult to understand for common genebank managers.

However, all participants were satisfied and appreciated the training course program. DOI is a new system to be implemented in rice germplasm management and exchange, and this was the first technical training course which provides very technical knowledge on how to assign and manage DOI to support the management and exchange of rice germplasm. Although some participants encountered problems during the training and did not successfully assign DOI to their rice accessions, they continued the registration after the training. Resource persons from ICABIOGRAD and IRRI continued to provide technical assistance to participants from all participating countries. Later, the Toolkit has been improved to overcome the complicated configuration and high resources requirement. The newest version (Toolkit v2.0.3) was released in 2018, and it had been successfully disseminated at the second DOI International Training Course in Yogyakarta, Indonesia from 26 April-3 May 2019. In this training course, 14 participants from 8 countries (Indonesia, Bangladesh, Pakistan, Myanmar, Cambodia, Lao PDR, Sri Lanka and Burundi) were present.

Registered DOI Achievement of Participating Countries

Currently, 8.419 accessions of PGRFAs from 13 countries have been assigned DOI. Detailed breakdown of the number of PGRFA accessions from each participating country is presented in Table 1 and Table 2. So far, Indonesia has registered as many as 1.369 PGRFA materials to obtain DOIs. Registration was done in 2017 (849 accessions), 2018 (351 accessions) and 2019 (169 accessions). In 2019, non-rice crops such as sweet potato, cassava, pigeon pea, banana, forages, citrus and maize were also included for the registration.

Since ICABIOGRAD is considered as the National Focal Point for PGRFA management in Indonesia, DOI registration of PGRFA collections from all domestic institutions is done under the coordination of ICABIOGRAD. Most of the PGRFA material is stored in the Genebank of ICABIOGRAD-IAARD, and some of them were contributed by other institutions under the IAARD, i.e.:

1. Assessment Institute for Agricultural Technology (AIAT) of Central Kalimantan (96 accessions of rice).
2. Assessment Institute for Agricultural Technology (AIAT) of East Kalimantan (10 accessions of rice).
3. Assessment Institute for Agricultural Technology (AIAT) of South Sumatera (55 accessions of rice).
4. Assessment Institute for Agricultural Technology (AIAT) of Yogyakarta (55 accessions of rice).
5. Assessment Institute for Agricultural Technology (AIAT) of West Nusa Tenggara (4 accessions of rice).
6. Indonesian Center for Rice Research (ICRR) (14 accessions of rice).
7. Indonesian Legumes and Tuber Crops Research Institute (ILETRI) (12 accessions of pigeon pea).
8. Indonesian Tropical Fruits Research Institute (ITFRI) (10 accessions of banana).
9. Indonesian Research Institute for Animal Production (IRIAP) (18 accessions of forages crops).
10. Indonesian Citrus and Subtropical Fruits Research Institute (ICSFRI) (12 accessions of citrus).
11. Indonesian Cereals Research Institute (ICERI) (20 accessions of maize).

Table 1. Registration status of participating countries after the first DOI training in 2018.

Country	WIEWS	Institution	Registered DOI		Crop
			During training	After training	
Indonesia	IDN179	Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD)	262	20	Rice
	IDN179	Assessment Institute for Agricultural Technology (AIAT) of Central Kalimantan	96	-	Rice
	IDN179	Assessment Institute for Agricultural Technology (AIAT) of South Sumatera	55	-	Rice
		Assessment Institute for Agricultural Technology (AIAT) of East Kalimantan	0	10	Rice
	IDN179	Assessment Institute for Agricultural Technology (AIAT) of Yogyakarta	0	55	Rice
		Assessment Institute for Agricultural Technology (AIAT) of West Nusa Tenggara	0	4	Rice
Zambia	ZMB048	National Plant Genetic Resources	60	269	Rice
Philippines	PHL158	Genetic Resources Division (GRD), Philippine Rice Research Institute (PhilRice)	0	1.016	Rice
India	IND001	Indian Council of Agricultural Research, National Bureau of Plant Genetic Resources	1.513	-	Rice
Malaysia	MYS005	Program of Genetic Resources and Germplasm Conservation Management (GB1) Genebank Seed and Centre, MARDI	0	708	Rice
Bhutan	BTN026	National Biodiversity Centre, Ministry of Agriculture and Forests	60	-	Rice
Total			1.895	2.082	

Table 2. Registration status of participating countries after the second DOI training in 2019.

Country	WIEWS	Institution	Registered DOI		Crop
			During training	After training	
Indonesia	IDN179	Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD)	5	30	Rice
			12	-	Sweet potato
			6	-	Casssava
	IDN179	Indonesian Legumes and Tuber Crops Research Institute (ILETRI)	12	-	Pigeon pea
	IDN179	Indonesian Tropical Fruits Research Institute (ITFRI)	10	-	Banana
	IDN179	Indonesian Research Institute for Animal Production (IRIAP)	18	-	Forages
	IDN179	Indonesian Citrus and Subtropical Fruits Research Institute (ICSFRI)	12	-	Citrus
	IDN179	Indonesian Cereals Research Institute (ICERI)	20	-	Maize
	IDN179	Indonesian Center for Rice Research (ICRR)	14	-	Rice
	IDN179	Agro-technology Innovation Center, Gadjah Mada University (AIC-GMU).	5 5 5 5 5 5	- - - - - -	Bean Winged bean Yard long bean Snow pea Eggplant Pumpkin
Bangladesh	BGD002	Genetic Resources and Seed Division, Bangladesh Rice Research Institute (BRRI)	70	-	Rice
Lao PDR	LAO018	Rice Research Center (RRC), National Agriculture and Forestry Research Institute (NAFRI)	47	393	Rice
Sri Lanka	LKA036	Plant Genetic Resources Center (PGRC)	94	3.402	Rice

Table 2. Continue.

Country	WIEWS	Institution	Registered DOI		Crop
			During training	After training	
Pakistan	PAK001	Plant Genetic Resources Program, Bio-resources Conservation Institute, National Agricultural Research Centre (NARC).	50	-	Rice
Cambodia	KHM010	Cambodian Agricultural Research and Development Institute (CARDI)	10	-	Rice
Burundi	BDI016	Burundi Genebank (BGB), Institut des Sciences Agronomiques du Burundi (ISABU)	188	-	Rice
Myanmar	MMR015	The Seed Bank of Myanmar Department of Agricultural Research (DAR)	49	-	Rice
Total			617	3,825	

ICABIOGRAD (Indonesia), PhilRice (Philippines), MARDI (Malaysia), RRC-NAFRI (Lao PDR) and PGRC (Sri Lanka) are the most active gene banks who continue to assign DOI to their PGRFA collection after the training courses. The PGRFA materials that have been registered so far from those gene banks are 1.369 accessions (ICABIOGRAD), 1.016 accessions (PhilRice), 708 accessions (MARDI), 440 accessions (RRC-NAFRI), and 3.496 accessions (PGRC). There are also 30 vegetable accessions (5 accessions each of bean, winged bean, yard long bean, snow pea, eggplant and pumpkin) being registered, which are managed by Agro-technology Innovation Center, Gadjah Mada University (AIC-GMU).

Concluding Remarks and Future Perspective

DOI system is very useful to be implemented in germplasm collection management and exchanges. Therefore, it should also be implemented for other crops in the gene bank collection. Some accessions of non-rice crops from ICABIOGRAD-IAARD Indonesian genebank had been successfully registered for DOI during the project implementation. The DOI system and the Toolkit have been disseminated to 13 participating countries from Asia and Africa. However, there is still a great chance to expand it to more participating countries in the future.

Acknowledgment

The authors thank the Benefit Sharing Fund of the ITPGRFA-FAO and European Union for financing the project.

REFERENCES

Alercia, A., López, F.M., Sackville Hamilton, N.R. & Marsella, M. (2018) Digital Object Identifiers for food crops-Descriptors and guidelines of the Global Information System. Rome: FAO.

Chandrakar, R. (2006) Digital object identifier system: an obverview. *The Electronic Library,* 24(4):445-452. Doi:10.1108/02640470610689151.

Information Standards Quarterly (ISQ) (2004) National Information Standards Organization (NISO). *Information Standards Quarterly*, 16(3):16.

Kurniawan, H., Kristiyono, E. & Hidayatun, N. (2019) Integration Toolkit v2.0.3: Installation and Operation Manual. The 2nd International Training on Software Testing to Support DOI Implementation on Rice Germplasm, Yogyakarta Indonesia 26 April – 2 May 2019. ICABIOGRAD, IAARD.

Paskin, N. (2010). Digital object identifier (DOI®) system. *Encyclopedia of library and information sciences,* 3:1586-1592. Doi:10.1081/E-ELIS3-120044418.

GENOMIC INFORMATION AS A POTENTIAL DESCRIPTOR OF GENETIC RESOURCES

Dani Satyawan

INTRODUCTION

Object identifier is an essential requirement in the management of genetic resources collection, as it provides information regarding the identity of each accession for both genebank managers and users. The digital object identifier (DOI) project aims to create a global standard in identifying plant genetic resources for agriculture (PGRFA) and provides a link between information that are available for a material and that material stored in a collection (Alercia et al. 2018). Currently, DOI contains several mandatory descriptors regarding taxonomic identity, date and method of procurement, as well as the location of the collection. Additional descriptors can also be added, but at the moment the suggested descriptors mostly contain geographical information.

While useful in multiple ways, geographical information quite often is just a snapshot of the current whereabouts of a PGRFA. This is because many PGRFA were introduced materials instead of native plants that evolve over a long period of time in a particular location. The absence of good record-keeping in many locations means that PGRFA exchange and introduction were often unnoticed. Thus, it is possible that materials collected from distant areas are actually closely related, which reduces their utility to be used as parental lines in breeding programs due to their low genetic diversity.

For breeding applications, geographical information can be useful in breeding programs for specific locations, such as drought-prone high altitude lands, swampy lands, and natural habitats of specific pathogens. However, geographical descriptor typically focuses on aiding in finding a location rather than providing agronomically relevant information about that location. Thus, breeders interested in materials adapted to certain environmental types will need to do more work to obtain the environmental characteristics of certain geographical locations.

Another notably useful type of information for breeders that are also absent in the descriptor is phenotypic information. This is understandable because phenotypic data is greatly influenced by the environment where they are scored. Thus, a good performance in one environment is not always a good predictor of similarly good performance in another location. This inconsistent nature makes it an unreliable identifier. The number of phenotypic characteristics that can be scored reliably is also limited, which means that a lot of PGRFA will have the same phenotypic descriptions. Hence, phenotypic data alone will not be able to function as unique identifiers for PGRFA.

On the other hand, genotypic data in the form of DNA sequence information can serve as a better identifier and descriptor. DNA sequence variation is the underlying cause of phenotypic variation, which is of great interest to plant breeders and farmers. Using recent advances in genomics research, it is possible to predict the phenotypes of a PGRFA at a reasonable accuracy using its DNA sequences (Hickey et al. 2017). The number of DNA variation in the genome of an organism is also more abundant, making it easier to find sequences that are unique and can be used as identifiers for individual PGRFA. The cost of DNA sequencing also tends to decrease while the genome's function becomes better characterized as time progresses (Van El et al. 2013), making DNA information more attractive as future descriptors and identifiers for PGRFA collection. This review discussed the capacity of current knowledge in genomics to be used as a descriptor and identifier of PGRFA, the technical challenges that need to be considered, as well as expected progress that can promote the utilization of genomic data in PGRFA management.

Current Capabilities of Genome Analysis

The cost of whole genome sequencing has gradually decreased and the choice of sequencing technology has expanded in recent years. As a result, many institutions have assembled high quality reference genomes for economically important PGRFA (table 1). High quality references are important because current sequencing technologies have not been able to read a complete plant genome uninterrupted from start to end. Reference genomes are usually assembled from overlapping short reads, which produces gaps in a large part of the genome, especially in highly repetitive regions in the chromosomes (Alkan, Sajjadian and Eichler 2011). Recent genome assemblies typically combined high quality short reads with long reads sequencing technologies to minimize missing sequences and better sequence fidelity across the whole genome (Shi et al. 2016).

Once a reference genome is assembled, the genome sequences of other accessions from the same species can be read more easily as their sequences can be deduced from their alignment with the reference genome. Consequently, many institutions have now sequenced a large number of PGRFA collections. One notable example is the 3000 rice genome project at the International Rice Research Institute, which sequenced 3048 rice accessions obtained from all over the world in IRRI genebank collection (Li, Wang and Zeigler 2014). The resulting sequence data was subsequently used to deduce the phylogenetic relationship among all of the accessions, and can be used as a context and reference when more rice accessions are sequenced and analyzed (Wang et al. 2018).

Table 1. List of crops with assembled reference genomes by 2013 (Michael and Jackson 2013).

Scientific name	Common name	Year	Chromosome (#)	Size (Mb)	Assembled (%)	Gene (#)	Journal	PMID
Oryza sativa	Rice	2005	12	389	95	37544	Nature	16100779
Populus trichocarpa	Black cottonwood	2006	19	485	84	45555	Science	16973872
Vitis vinifera	Grape	2007	19	475	103	30434	Nature	17721507
Carica papaya	Papaya	2008	9	372	99	28629	Nature	18432245
Sorghum bicolor	Sorghum	2009	10	818	90	34496	Nature	19189423
Zea mayes	Maize	2009	10	2300	89	32540	Science	19965430
Glycine max	Soybean	2010	20	1115	87	46430	Nature	20075913
Ricinus communis	Castor bean	2010	10	320	102	31237	Nature Biotechnology	20729833
Malus x domestica	Apple	2010	17	742	81	57386	Nature Genetics	20802477
Jatropha curcas	Jatropha	2010	NA	380	75	40929	DNA Research	21149391
Theobroma cacao	Cocoa	2011	10	430	76	28798	Nature Genetics	21186351
Fragaria vesca	Strawberry	2011	7	240	87	34809	Nature Genetics	21186353
Phoenix dactylifera	Date palm	2011	18	658	58	28890	Nature Biotechnology	21623354
Solanum tuberosum	Potato	2011	12	844	86	39031	Nature	21743474
Cucumis sativus	Cucumber	2011	7	367	88	26587	PlosOne	21829493
Brassica rapa	Chinese cabbage	2011	10	485	59	41174	Nature Genetics	21873998
Cannabis sativa	Hemp	2011	?	820	96	30074	Genome Biology	22014239
Cajanus cajan	Pigeon pea	2011	11	833	72	48680	Nature Biotechnology	22057054
Solanum lycopersicum	Tomato	2012	12	900	84	34727	Nature	22660326
Cucumis melo	Melon	2012	12	450	83	27427	PNAS	22753475

Table 1. Continue.

Scientific name	Common name	Year	Chromosome (#)	Size (Mb)	Assembled (%)	Gene (#)	Journal	PMID
Linum usitatissimum	Flax	2012	15	373	85	43484	Plant Journal	22757964
Musa acuminata malaccensis	Banana	2012	11	523	90	36542	Nature	22801500
Gossypium raimondii	Cotton D	2012	13	880	88	40976	Nature Genetics	22922876
Azadirachta indica	Neem	2012	NA	364	NA	20169	BMC Genomics	22958331
Hordeum vulgare	Barley	2012	7	5100	98	30400	Nature	23075845
Citrullus lanatus	Watermelon	2012	11	425	83	23440	Nature Genetics	23179023
Triticum aestivum	Wheat	2012	21	17000	22	94000	Nature	23192148
Gossypium raimondii	Cotton D	2012	13	880	84	37505	Nature	23257886
Prunus mume	Chinese plum	2012	8	280	85	31390	Nature Communication	23271652
Pyrus bretschneideri	Pear	2013	17	527	97	42812	Genome Research	23149293
Cicer arietinum	Chickpea	2013	8	738	72	28269	Nature Biotechnology	23354103
Hevea brasiliensis	Rubber tree	2013	18	2150	52	68955	BMC Genomics	23375136
Prunus persica	Peach	2013	8	265	86	27852	Nature Genetics	23525075
Aegilops tauschii	Wheat DD	2013	7	4360	97	43150	Nature	23535592
Triticum urartu	Wheat AA	2013	7	4940	94	34879	Nature	23535596

Using association analysis, DNA segments that correlate with several important traits were identified (Huang et al. 2012). In the future, as this kind of information and the results of individual gene function studies accumulate, they can be used as the basis to predict traits using DNA sequence data. Many research groups are currently developing prediction algorithms that utilize such data in conjunction with DNA variation data (Spindel et al. 2016). Some of those predictive algorithms were even deemed good enough to be used in breeding programs to aid or even replace physical evaluation in the field (Heffner et al. 2010). Breeders and farmers who try to find plant accessions that carry certain traits will be able to use such algorithms to select potential parental candidates among thousands of materials stored in genebanks.

DNA sequences also carry information that can be used to identify the organism that carry the DNA (Čandek and Kuntner 2015). Several toolkits were developed to identify the taxonomic identity of an organism based on the DNA sequence of parts of its genome (Liu et al. 2017). The availability of whole genome sequences enable this analysis to not only identify which species the sample belongs to, it even indicates the population and families the sample is most closely related to. Using whole genome sequence data from more than 3000 well-characterized rice accessions from IRRI as reference points, it is now possible to predict the origin and subpopulation of recently sequenced rice accessions, as well as elucidate their pattern of genetic admixture and ancestry (Wang et al. 2018). This will be useful when intellectual property dispute arises from the utilization of genetic resources.

DNA sequence data can also be used to overcome a common problem experienced by genebanks, namely the presence of redundant and duplicated samples. Such problem arose because plant variety exchange is relatively common across many regions, which means that several plants cultivated in different areas under different cultivar names are actually the same plant material. Lack of proper documentation system and sample maintenance management also often cause mislabeling of genebank accessions. Duplicated samples increase the cost of genetic resource maintenance while offering little value for those interested in utilizing genebank collection. DNA sequence data can easily identify and eliminate such accessions, thus improving the efficiency of genebank management (Singh et al. 2019).

Challenges of Incorporating Genomic Information

Although genomic information has several intrinsic advantages, its application as descriptors or passport data presents some new challenges for the systems that are currently used in most genebanks. For a start, the size of genome information is significantly larger than the data currently inputted as descriptors in genebank database system. Such large information necessitates the use of specialized computers, network, and software to ensure smooth and immediate access of the DNA information. All of those equipments can represent a significant investment that is out of reach for most genebanks. Fortunately, specialized cloud computing platforms and network providers that can handle such demand is becoming more ubiquitous and hopefully will become more economically viable for genebanks in the immediate future (Langmead and Nellore 2018).

Alternatively, genomic DNA variation data can be compressed by focusing only on DNA sequences that are different from the reference genome sequence. Trait variations arise from DNA sequences that are different between each accession, and even in different species like humans and bonobo DNA sequence variations only account for less than 2% of the genome (Prüfer et al. 2012). Thus, a large portion of the genomic data that are identical in all members of the species can be excluded to reduce the size of incorporated genomic DNA data. However, exclusive use of variable DNA sequences also necessitates the inclusion of sequence coordinate data to provide information regarding the location of the DNA variation. Those coordinate data can also significantly increase file size and add another type of data that computers need to read and decipher.

Another potential hurdle is that DNA information is relatively opaque and specialized software is required to translate DNA sequence data into more useful information, such as sample identity and potential ability to express useful traits. Currently, there is no single software that can provide all of that information from DNA sequence data, as most available softwares were designed to do specific types of sequence analysis. Genebanks intending to provide DNA analysis services therefore must also provide various DNA analysis programs and provide support to their users on how to operate that software. This can create a significant burden as some of those software also require specific training to operate and specific computer hardware to run (Schneider et al. 2010).

Potential users with no background knowledge in bioinformatics may also feel discouraged when faced with complex software and analysis pipelines. Current software that can predict traits from DNA sequence data also cannot do so in a highly accurate manner (Spindel et al. 2015). Although it can still be useful in some applications, the inaccuracy can be problematic for some crops that are expensive to maintain as every mistake represents a significant investment loss for the recipient. However, as the function of each part of the genome is gradually being elucidated, it can be expected that the prediction accuracy will improve in the future. This of course assumes that such studies will be funded in the future, while in reality not all crop species are allocated sufficient funding for even basic genetic research. A lot of less popular tropical crops do not yet have proper reference genomes and have rarely been the subject of genetic studies. For such orphaned crops, DNA information usefulness may come further away in the future.

Expected Progress in the Field of Genomics

It is clear that a reference genome is needed for maximum utilization of genomic information. The good news is that aside from the ever-decreasing sequencing cost, new sequencing technologies have also been developed which greatly assist in reference genome assembly, such as long read sequencing technologies and optical mapping (Dong et al. 2013; Rhoads and Au 2015). Higher quality reference genomes had been generated using combinations of sequencing technologies at a quicker pace and lower cost, which improves the chance that even orphaned crops will soon have their own high quality reference genomes. This had been initiated in the 1000 plant genome project, which sequenced only the genes or expressed portion of all DNA in the genome of 1000 plant species (Matasci et al. 2014). An even more ambitious project to sequence the full genome of 10.000 plant species, protists, and algae is also underway and is expected to be completed by the year 2023 (Cheng et al. 2018).

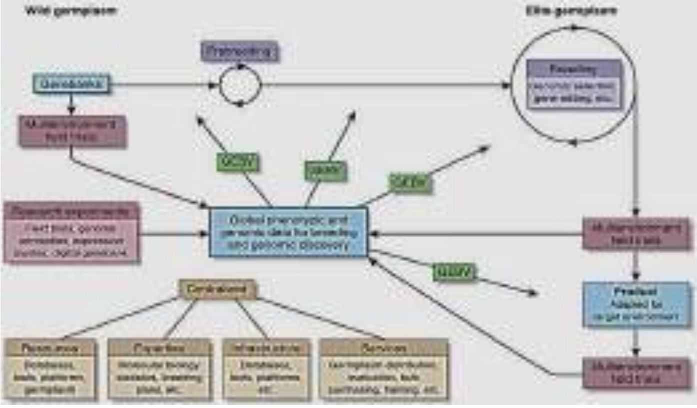


Figure 1. Proposed global framework for combining animal and plant breeding approaches to accelerate genetic gain through coordinated phenotyping, sequencing, genomic prediction, and genome editing (Hickey et al. 2017).

Those reference genomes will need to be accompanied by comprehensive annotation regarding the function of each segment of the genome. Thus, some forms of automation may be needed to complete such a massive undertaking (Cheng et al. 2018). The result may be rudimentary, but it will be a good foundation for researchers to perform genetic analysis later on. The inter-relatedness of all organisms on earth and the fact that DNA code is relatively universal also mean that data from one species sometimes translate well when applied to closely related species (Proost et al. 2009). Thus, software that predict trait from DNA information could be applicable not only to the species it was originally designed for, but also for other species related to that species.

On the software side, improvements need to be made on the usability, accuracy, and integration aspect. Hickey et al. (2017) proposed a global framework to integrate phenotypic and genomic data generated from coordinated research experiments to increase genebank utilization and incorporate genomic prediction into plant breeding programs (Figure 1). The approach also integrate best practices from animal breeding, such as selection based on genomic estimated breeding values (GEBV), to accelerate product generation in plant breeding. The collaborative aspect will also encourage standardization and consistency, which contrasts to the highly fragmented and specialized software development found in today’s environment.

Lastly, the advent of genome editing technology will enable plant improvement that is based on genomic information without the need for physical genetic resources (Roa et al. 2016). When such technology becomes easy to implement, transfer of genetic information becomes more important than the genetic materials themselves. Consequently, genebanks may no longer need to keep as many collection of plant materials as they do today and replace some of the collections with their genetic information instead. The use and management of DNA information are therefore inevitable and the sooner we figure out how to do it efficiently, the smoother the transition will be.

Conclusion

The use of genomic DNA information as one of the descriptors for genetic resources carries some intrinsic advantages that are beneficial for PGRFA management and utilization. Although currently available technology and infrastructure may prevent its effective implementation immediately, research progress in the field of computing and genomics can speed up its feasibility. Genomic information gradually becomes an important element in modern breeding programs and may eventually eliminate the need for seeds to be used as parental materials when genome editing technique becomes sufficiently advanced for regular applications. Consequently, genebanks need to start incorporating DNA information in their workflow to stay relevant and the formulation and adoption of DOI can be used as a starting point to develop such a system.

REFERENCES

Alercia, A. et al. (2018) Digital object identifiers for food crops: descriptors and guidelines of the global information system. FAO. Available from http://agris.fao.org/agris-search/search.do?recordID=XF2018002129.

Alkan, C., Sajjadian, S. & Eichler, E.E. (2011) Limitations of next-generation genome sequence assembly. *Nature Methods*, 8(1):61-65. Doi:10.1038/nmeth.1527.

Čandek, K. & Kuntner, M. (2015) DNA barcoding gap: reliable species identification over morphological and geographical scales. *Molecular Ecology Resources,* 15(2):268-277. Doi:10.1111/1755-0998.12304.

Cheng, S. et al. (2018) 10KP: A phylodiverse genome sequencing plan. *GigaScience*, 7(3). Doi:10.1093/gigascience/giy013.

Dong, Y. et al. (2013) Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nature Biotechnology*, 31(2):135-141. Doi:10.1038/nbt.2478.

van El, C.G. et al. (2013) Whole-genome sequencing in health care. *European Journal of Human Genetics,* 21(6):580-584. Doi:10.1038/ejhg.2013.46.

Heffner, E.L. et al. (2010) Plant breeding with genomic selection: gain per unit time and cost. *Crop Science,* 50(5):1681-1690. Doi:10.2135/cropsci2009.11.0662.

Hickey, J.M. et al. (2017) Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature Genetics,* 49(9):1297-1303. Doi:10.1038/ng.3920.

Huang, X. et al. (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics,* 44(1):32-39. Doi:10.1038/ng.1018.

Langmead, B. & Nellore, A. (2018) Cloud computing for genomic data analysis and collaboration. *Nature Reviews Genetics,* 19(4):208-219. Doi:10.1038/nrg.2017.113.

Alercia, A. et al. (2018) Digital object identifiers for food crops: descriptors and guidelines of the global information system. FAO. Available from http://agris.fao.org/agris-search/search.do?recordID=XF2018002129.

Alkan, C., Sajjadian, S. & Eichler, E.E. (2011) Limitations of next-generation genome sequence assembly. *Nature Methods,* 8(1):61-65. Doi:10.1038/nmeth.1527.

Čandek, K. & Kuntner, M. (2015) DNA barcoding gap: reliable species identification over morphological and geographical scales. *Molecular Ecology Resources,* 15(2):268-277. Doi:10.1111/1755-0998.12304.

Cheng, S. et al. (2018) 10KP: A phylodiverse genome sequencing plan. *GigaScience*, 7(3). Doi:10.1093/gigascience/giy013.

Dong, Y. et al. (2013) Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nature Biotechnology,* 31(2):135-141. Doi:10.1038/nbt.2478.

van El, C.G. et al. (2013) Whole-genome sequencing in health care. *European Journal of Human Genetics,* 21(6):580-584. Doi:10.1038/ejhg.2013.46.

Heffner, E.L. et al. (2010) Plant breeding with genomic selection: gain per unit time and cost. *Crop Science,* 50(5):1681-1690. Doi:10.2135/cropsci2009.11.0662.

Hickey, J.M. et al. (2017) Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature Genetics*, 49(9):1297-1303. Doi:10.1038/ng.3920.

Huang, X. et al. (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics,* 44(1):32-39. Doi:10.1038/ng.1018.

Langmead, B. & Nellore, A. (2018) Cloud computing for genomic data analysis and collaboration. *Nature Reviews Genetics,* 19(4):208-219. Doi:10.1038/nrg.2017.113.

Li, J.-Y., Wang, J. & Zeigler, R.S. (2014) The 3,000 rice genomes project: new opportunities and challenges for future rice research. *GigaScience,* 3(1):8. Doi:10.1186/2047-217X-3-8.

Liu, J. et al. (2017) Multilocus DNA barcoding – Species Identification with Multilocus Data. *Scientific Reports,* 7(1):16601. Doi:10.1038/s41598-017-16920-2.

Matasci, N. et al. (2014) Data access for the 1,000 Plants (1KP) project. *GigaScience*, 3(1):17. Doi:10.1186/2047-217X-3-17.

Michael, T.P. & Jackson, S. (2013) The First 50 Plant Genomes. *The Plant Genome,* 6(2). Doi:10.3835/plantgenome2013.03.0001in.

Proost, S. et al. (2009) PLAZA: a comparative genomics resource to study gene and genome evolution in plants. *The Plant cell,* 21(12):3718-31. Doi:10.1105/tpc.109.071506.

Prüfer, K. et al. (2012) The bonobo genome compared with the chimpanzee and human genomes. *Nature,* 486(7404):527-531. Doi: 10.1038/nature11128.

Rhoads, A. & Au, K.F. (2015) PacBio sequencing and its applications. *Genomics, Proteomics & Bioinformatics,* 13(5):278-289. Doi:10.1016/J.GPB.2015.08.002.

Roa, C. et al. (2016) Plant genetic resources: needs, rights, and opportunities. *Trends in Plant Science*. 21(8):633-636. Doi:10.1016/J.TPLANTS.2016.06.002.

Schneider, M.V. et al. (2010) Bioinformatics training: a review of challenges, actions and support requirements. *Briefings in Bioinformatics*, 11(6):544-551. Doi:10.1093/bib/bbq021.

Shi, L. et al. (2016) Long-read sequencing and de novo assembly of a Chinese genome. *Nature Communications,* 7(1):12065. Doi:10.1038/ncomms12065.

Singh, N. et al. (2019) Efficient curation of genebanks using next generation sequencing reveals substantial duplication of germplasm accessions. *Scientific Reports,* 9(1):650. Doi:10.1038/s41598-018-37269-0.

Spindel, J. et al. (2015) Genomic selection and association mapping in rice (*Oryza sativa*): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. Mauricio, R. (ed.) *PLOS Genetics*, 11(2):e1004982. Doi:10.1371/ journal.pgen.1004982.

Spindel, J.E. et al. (2016) Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. *Heredity,* 116(4):395-408. Doi:10.1038/hdy.2015.113.

Wang, W. et al. (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature,* 557(7703):43-49. Doi:10.1038/s41586-018-0063-9.

CURRENT AND FUTURE OF DOI IMPLEMENTATION IN THE LAO NATIONAL GENEBANK

Koukham Vilayheuang, Sipharphone Aloun,
Chanthakhone Boualaphanh, and Chay Bounphanousay

INTRODUCTION

Lao national genebank (LNGB) is located in Phonthong village, Xaythany district, Vientiane capital, Lao People Democratic Republic (Lao PDR). LNGB was established in 1995 under The Rice Research Center of National Agriculture and Forestry Research Institute (NAFRI). It is the main genebank in Lao People's Democratic Republic (Lao PDR) which has function for seeds conservation in certain conditions. Recently, there are many crops conserving inside, such as rice, maize, sorghum and millet. Among those mentioned crops germplasm, rice is the biggest number as 14.000 accessions conserved in the genebank. Rice is a staple food of mostly peoples in the world (Mohanty 2013), and most people in Lao depend on it as their staple food and rice dominates agricultural production. Since 1995, NAFRI and International Rice Research Institute (IRRI) conserved all of rice accessions throughout the country.

During the collection, the genebank provided many accessions of rice for users. However, it is difficult to track the provided samples from users. Therefore the Digital Object Identifier (DOI) now has been chosen as problem solver to provide a globally unique and permanent mechanism to track and identification of germplasms. DOI is a standardized alphanumeric string that is assigned by a registration agency and provides a persistent link to the location of information about the object on the internet (Paskin 2010). Digital object identifier (DOI) was first introduced to LNGB manager through the DOI Implementation training in Bali, Indonesia 2019. In this paper, we will describe the implementation of DOI in Lao National Genebank.

Collection and Conservation of PGRFA in the Lao PDR

Rice germplasm in Lao was started to explore for conservation and utilization in 1970s. Another 1000 rice germplasm were collected from 1991 to 1994 (Roder et al. 1996). IRRI regarded a set of 13.192 samples collected from 1995 to 2000 as being representative of the PGRFA throughout the country (Rao et al. 1996). However there was no genebank available in the country to conserve those genetic materials. National institutions of Lao in collaboration with international organizations have taken measures to identify and conserve valuable genetic material for future use. With the aid of international funding organization, ARC under NAFRI has initiated biodiversity programmes for several crops including rice. There is a joint project namely "Rice Biodiversity Project for Rice Germplasm Collection" between Lao's Ministry of Agriculture and IRRI called the Lao-IRRI project until 1993 (Thepphavong & Sipaseuth 2007).

The objectives of the Lao components of Lao-IRRI project were collect traditional rice varieties cultivated under diverse agroclimate conditions and under different ecosystems throughout the country, collect wild and spontaneous interspesific hybrids between wild and cultivated forms of rice, preserve the collected wild and cultivated forms in appropriate storage facilities at NARCin Sythany Distric of Vientiane Municipality, conserve a duplicate set of the germplasm collection at the International Rice Genebank (IRG) at IRRI, and make the germplasm collection available for utilization in Lao PDR and elsewhere (Bounphanousay et al. 2000). The National Genebank was established in 1995. Collection of rice germplasm started in 1993 and lasted till 2000, and approximately 14.000 samples of rice germplasm were collected.

Distribution of Rice Germplasm to Users

There is big collection of rice germplasm throughout the country during 1993 to 2000, and the rice germplasm are not only conserved in the LNGB, but also in IRRI (Rao et al. 2006). Some of rice germplasms are conserved in the LNGB for future purposes, and some of them were provided to the users for research purpose (Table 1).

Current DOI Implementation

The current progress of DOI Implementation in the LNGB is there are 440 rice accessions were registered and obtained with DOI number (https://ssl.fao.org/glis/entity/search) since the genebank manager of LNGB was participated in DOI training in Bali, Indonesia in 2019. However, compare to the whole rice germplasm collection in the LNGB Genebank, only 2.9% accession registered in DOI system and 97.1% have not yet assign to DOI.

Table 1. Distribution of rice germplasm in the LNGB.

Users	Germplasm requested	Reference
1	Rice	Bounphanousay et al. 2008
2	Rice	Kanyavong 2012
3	Rice	Boualaphanh et al. 2011
4	Rice	Vilayheuang et al. 2016

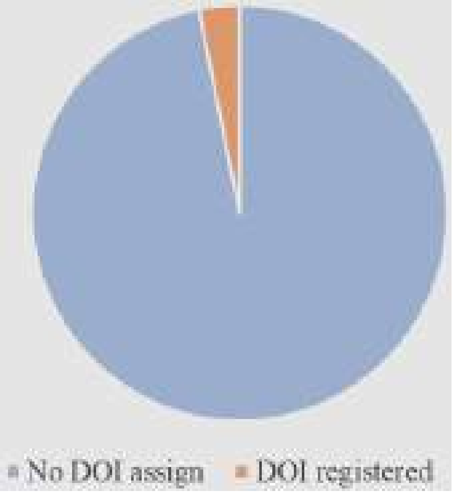


Figure 1. Showing the ratio between rice samples DOI registered and non-DOI.

Concluding Remarks and Future Perspective

In Lao PDR, most of food crops germplasm collections are conserved in the National Genebank. However, some of them were distributed to users and the germplasm tracking system was not yet available. Recently, 440 rice accessions were registered for DOI which could enable genebank to tracking on it. Even though, only few percent of number of accession registered compare to the whole collection, but it presenting the beginning of DOI implementation in LNGB. In the future, all accessions of rice and other crops are planning to be registered. Thus, capacity building programmes both national and international for LNGB staffs should be considered in order to increase the implementation of DOI in LNGB.

Acknowledgement

We acknowledged the Indonesian Center for Agricultural Biotechnology and Genetic Resource Research and Development (ICABIOGRAD) for giving us opportunity to participate in the DOI training.

REFERENCES

Bounphanousay, C., S. Appa Rao, K. Kanyavong, B. Sengthong, J.M. Schiller and M.T. Jackson. (2000) Conservation of Lao Rice Germplasm at the International Rice Genebank, IRRI, Philippines. Part 5, Collection Period: October 1999 to March 2000. Vientiane, Lao-IRRI Project. 44 p.

Boualaphanh, C., Calingacion, M., Cuevas, R. P., Jothityangkoon, D., Satichon, J. & Fitzgerald, M. (2011) Yield and quality of traditional and improved Lao varieties of rice, *Science Asia*, 37: 89-97. Doi: 10.2306/scienceasia1513-1874.2011.37.089.

Bounphanousay, C., Jaisil, P., McNally, K.L., Sanitchon, J., Sackville Hamilton, N.R. (2008) Variation of microsatellite markers in collection of Lao's black glutinous rice (*Oryza sativa* L.). *The Lao Journal of Agriculture and Forestry*, 18:32-47.

Kanyavong, K. (2012) Study on genetic diversity of khao Kai Noi (small chiken rice) local lao rice germplasm by morphological characters. MSc (Agronomy) theses. National University of Laos, Lao PDR.

Mohanty, S. (2019) Trends in global rice consumption. *Rice Today Newsl,*12:44-45. Available from: http://www.scribd.com/doc/119860372/RT-Vol-12-No-1-rice-facts.

Paskin, N. (2010) Digital object identifier system. *Encyclopedia of Library and Information Sciences, Third Edition.* Doi: 10.1081/E-ELIS3-120044418.

Rao, S.A., Schiller, J.M., Bounphanousay, C. & Jackson, M.T. (2006) Diversity within the traditional rice varieties of Laos. *In*: Schiller, J.M., Chanphengxay, M.B., Linguist, B. & Rao, S.A. (eds) Rice in Laos, Los Baňos, Philippines, IRRI, pp. 123-140.

Roder, W., Keoboulapha, B., Vannalath, K. & Phouaravanh (1996) Glutinous rice and its importance for hill farmers in Lao. Econ. Bot. 50:401-408.

Thepphavong, B. & Sipaseuth (2007) Report of plant breeding and related biotechnology capacity: Lao people's democratic republic. *Global Partnership Initiative for Plant Breeding Capacity Building.*

Vilayheuang, K., Ryoko-Machida, H. Bounphanousay, C. & Watanabe, N.K. (2016) Genetic diversity and population structure of 'Khao Kai Noi', a Lao rice (*Oryza sativa* L.) landrace, revealed by microsatellite DNA markers. *Breeding Sciences*, 66:204-212 Available at: https://ssl.fao.org/glis/entity/search [Accessed July 20, 2019].

DIGITAL OBJECT IDENTIFIERS FOR EASY ACCESS OF INFORMATION AND EXCHANGE OF GERMPLASM

Min San Thein

INTRODUCTION

Myanmar is one of the centers of genetic diversity of rice and varietal group which spreads along the Himalayas from Iran to Myanmar consists of very diverse varieties (Glaszmann 1986). Some plants species exist in different pytopeographic regions with great antiquity and famous of its richness. However loss of crop genetic diversity for many reasons in the country has warned to conserve them in proper ways. At present more than 12000 accessions of different crop germplasms are being conserved in Myanmar Seed Bank. The diverse germplasm collection, however, are challenged for crop improvement through breeding programs for desirable traits in national level. Therefore exchange of germplasm and related information become crucial, and Digital Object Identifiers (DOI) can play a key role in this issue. By using the DOI standards, users will be able to identify and document their plant material uniquely and permanently, and will facilitate data interoperability among different systems (Alercia et al. 2018).

Germplasm Information in Myanmar Seed Bank

Myanmar Seed Bank (MSB) was established in 1990 by a grant of JICA of Japanese government and Myanmar government. During 1990-1992, a number of scientists from international organizations visited Myanmar to discuss further to explore for he compete of germplasn collection, conservation and management of plant genetic resources. Systematic genebank activities such as exploration, characterization, cold storage and database were started in 1997. Collections of landraces for different crops are used in collaboration of international organizations. Currently, a total of 8056 rice accessions along with other crop genetic resources (Table 1) are being conserved in *ex situ* collection in MSB (Annual Report 2019). At the beginning, information on data passport and characterization data is available only in hard copy catalogues. These limitations suggest a better management in the next period.

In 2009 the FAO Regional Project (GCP/RAS/240/JPN) has been designed for Strengthening Information on Plant Genetic Resources in Asia including MSB, called as National Information Sharing Mechanism Global Plan of Action (NISM-GPA). The website of NISM-GPA provides limited information related to plant genetic resources since they were not a big digitised data set like data passport and characterization data. Since Myanmar is a member of International Treaty on Plant Genetic Resources for Food and Agricultural (ITPGRFA), its participation should be inclusion of multilateral system by exchanging information and material of germplasm for food and agriculture. In this regard, huge data sets of rice and other crop germplasms conserved in the MSB should be available in the network of global genebanks. Collaboration between Myanmar with neighboring countries and developed countries are also important to achieve food security.

Table 1. Number of *ex situ* germplasm accessions in Myanmar Seed Bank (2019).

Sr. no.	Crop group	Number of accessions
1	Rice	8056
2	Legume	1578
3	Cereal	2238
4	Oilseed	800
5	Others	166
Total		12838

Training on Software Testing to Support DOI Implementation

To increase the capacitiy building, Myanmar participated in the 2nd International Training on Software Testing to Support DOI Implementation which was held in Yogyakarta, Indonesia during 26 April-3 May 2019. The training course lasted one week and was attended by genebank managers, database specialists and instructors from 8 countries (Figure 1). The participants have learned genetic resources related law, treaty such as ITPGRFA and Nagoya Protocol and have practiced installation of global information system (GLIS) toolkit v.2.0.3, operation of manual, input of huge data sets and software testing, finally have studied the implementation of DOIs then the registration of the data for exchanging information and germplasm material among genebanks. This training is beneficial for Myanmar particularly to support genetic resources management which could be involve in global information system.

Figure 1. Participants of the 2nd International Training on Software Testing to Support DOI Implementation on Rice Germplasm.

Global Information System for Germplasm Utilization

In the past, crop germplasms was managed at the genebank. Over the years, other systems such as EURISCO, SINGER and GRIN have been developed and integrated databases from various genebanks together with powerful tools for accessing and searching germplasm accessions information (Sebastian and Mackay 2009). At present, Global Information System (GLIS) comes into force of the ITPGRFA and it will contribute to access and benefit sharing of such information available to all contracting parties of the Treaty that can be as a user-friendly internet portal/gateway to information at the accession level (Quek 2009). Thus, near the future germplasm material and its related information will be easily accessed by the stakeholders of genetic resources.

Acknowledgment

The author would like to thanks the Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD), for inviting to participate in the 2nd International Training on Software Testing to Support DOI Implementation on Rice Germplasm that was held in Yogyakarta, Indonesia on 26 April – 3 May 2019.

REFERENCES

Annual Report (2019) Seed Bank Annual Report, Department of Agricultural Research, Yezin, Nay Pyi Taw, 2019 March.

Alercia, A., Lopez, F.M., Sackville Hamilton, N.R. & Marsella, M. (2018) Digital Object Identifiers for food crops. Descriptors and guidelines of the Global Information System. Rome: FAO.

FAO (2009) Strengthening information on plant genetic resources in Asia. Bangkok, Thailand: FAO Regional Office.

Glaszmann, J.C. (1986) A varietal classification of Asian cultivated rice (*Oryza sativa* L.) based on isozyme polymorphism. In "Rice Genetics" Proceedings of the International Rice Genetics Symposium. Manila, Philippines: IRRI.

Quek, P. (2009) Experiences with NISM – GPA and Complementary PGR Information Systems: Lecture Handout in Training for Trainers on The National Infromation Sharing Mechanism Global Plan of Action (NISM – GPA), Bangkok, 2009.

Sebastian, L. & Mackay M. (2009) Plant Genetic Resources Information System of Bioversity International and CG System (Update).

ROLE OF DIGITAL OBJECT IDENTIFIERS FOR NETWORKING AND MANAGING DATA TO AGRICULTURE BIODIVERSITY'S FUTURE

Bigirimana Jean Claude and Mbonihankuye Cyrille

INTRODUCTION

Burundi gene bank (BGB) was established in 2004, thus prior activities have been devoted on bringing together resources and skills needed for the genebank management. The main objectives of BGB are (1) to collect and conserve plant genetic resources of traditionally grown; (2) to strengthen scientific and technical capacities for the sustainable management of plant genetic diversity; (3) to establish and improve the use of data on plant genetic resources in Burundi; (4) to raise a public authorities' awareness on the importance on these resources in agricultural research plans to the national level and worldwide (ISABU 2009).

At more than one decade, data on Burundi plant genetic resources for food and agriculture (PGRFA) remain inaccessible for use by researchers and breeding programs worldwide. This is inconsistent with the objectives of the International Treaty on

Plant Genetic Resources for Food and Agriculture (ITPGRFA) on the use of genetic resources and the fair and equitable sharing of benefits derived from their use, in harmony with the Convention on Biological Diversity (CBD) for sustainable agriculture and food security. The challenge of BGB was not to be able for long time to share information to potential users on the actuality of PGRFA diversity. Even though plant materials were conserved, rational use has not been much facilitated. Due to lack of a reliable system for sharing descriptive information on biodiversity, information of PGRFA from scientific users has not been collected properly

International Cooperation on Plant Conservation

The Institut des Sciences Agronomiques du Burundi (ISABU) built a Genebank of Burundi (BGB) as recommended by the Government. It works in close collaboration with scientific and technical partners facilitated by international agencies which are involved in agricultural research.

At the beginning, funding was jointly provided by the Government for long-term and by the Swedish International Development Agency (SIDA) in short term. International funding was executed through EAPGREN (East African Plant Genetic Resources Network) covering Burundi, Eritrea, Ethiopia, Kenya, Madagascar, Rwanda, Sudan and Uganda. The collaboration work were with Nordgen (Nordic Genetic Bank-Sweden) which strengthened scientifically and technically network; Bioversity International; and Global Crop Diversity Trust (Bashaasha et al. 2010). BGB also collaborated with universities and multidisciplinary organizations involved in protection of the environment in national level. Importantly, Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and

Development (ICABIOGRAD) invited BGB to participate in the training for the Digital Object Identifier (DOIs) assignation to rice germplasm. Through this training, Burundi has been suggested to assign DOIs for plant accession to be accessed on global scale through FAO's Global Information System (GLIS) (https://ssl.fao.org/glis). This web portal provides user-friendly tools to develop a dynamic system of exchange and access to information. It helps researchers, breeders, farmers and other users to access the relevant scientific and technical information.

Profile of PGRFA Collections Conserved in the BGB

Since the BGB inception in 2004, there is significant progress in the inventory and collection of accessions of traditional food crop species across the country. These activities resulted in collection of 2044 accessions grouped in 43 species including 987 accessions of beans and 229 of sorghum in which the two crops have large traditional gene pools in Burundi. There are also 219 accessions of rice, 130 of maize, 143 of cassava, 204 of banana, 57 of groundnut, 67 of finger millet and 326 accessions for other species especially vegetables (Figure 1). Grains are then stored in freezers at-18°C (ISABU, 2018).

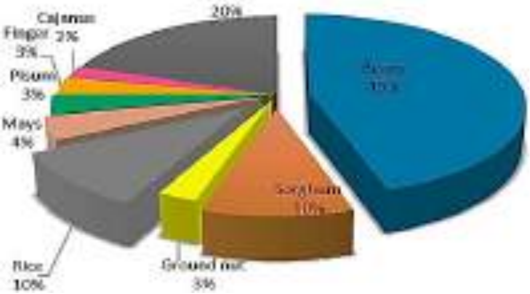


Figure 1. Proportions of seeds of several plant species stored in the Burundi Genebank

Germplasm accessions are duplicated and hermetically packaged in aluminium bags to be stored as active collections accessible at any time for studies and used for breeding programs. Copies of accessions have been sent to international gene banks as a security collection for long-term use. Burundi PGRFA collections are conserved in two international gene banks i.e. 439 accessions of 11 crops are kept at Svalbard Global Seed Vault (SGSV-Norway) and 283 accessions of beans are deposed at IITA (Nigeria). As many as 124 accessions of rice have been sent to Africa rice for molecular characterization (ISABU 2018). Diversity of vegetative propagated cultivars such as cassava, yam and taro are conserved on farm. These crops are used for food security and resilience to climate change as a result of their ability to drought adaptation (Bigirimana et al. 2016). Data of all these accessions have been recorded for long time and saved in Excel format and also identified by geographic coordinates.

Importance of the DOI Registration on Burundi Gene Bank: Case Study on Rice Germplasm

Burundi faced a problem on data documentation system. The lack of reliable digit data challenged the BGB for recording characterization in the safe way. The recording data with DOI could contribute to track and update data base on genetic resources and be consulted by scientific users around the world. Digitalization of genetic resources information is particularly an answer to the BGB gap.

Burundi's participation to training on DOI implementation in Indonesia in 2019 has enhanced the capacity building of its staff for improving and adjust database to provide comprehensive information to users. Registration on DOI addresses issues related to informations needed by curators, breeders for their breeding programs, and researchers to support germplasm conservation and utilization. Sharing data from BGB with the Global Information System facilitates easy access of informations related to crops diversity for research activities and breeding programs. The digitalization database could be beneficial for:

- Gene bank managers: to evaluate and quantify the degree of variability among accessions in gene bank;
- Breeders: to grasp and use existing diversity to develop more effective breeding strategies and achieve the goals of producing high yield to feed current and future generations;
- The government: develop in partnership with farmers and other managers of plant genetic resources conservation priorities to maintain the diversity.

The training DOI implementation introduced the digital system for rice germplasm management for conservation and exchange purposes. Testing the software to support the implementation of DOI is focude on rice because this plant species is one of most important crop widespread around the world. The key topics developed were: (1) software installation and configuration of the toolbox; (2) data setting in conformity with required standard compatible with the toolkit; (3) Importing data from the adapted file; (4) application of toolkit V2.0.3 (on a GLIS test server); (5) Storing DOI Recording Results; (6) Error handling after registration (GLIS 2019). The training allowed each participant to establish and test the software through humble manipulations, improve database and DOI registration of accessions. Training manuals and GLIS toolkit V.2.0.3 for guidance in setting up and operationalization of the software were provided.

Progress on Documentation and Data Sharing on Burundi PGRFA

During 12 years of data recording, BGB had not yet found any electronic documentation network to share and consult data (ISABU 2018). Geographical coordinates and identity of preserved accessions were not available for various scientific uses and for agricultural development purposes. Data have been for long time non-exchangeable efficiently between BGB with any regional or international gene banks. The use of data sharing tool became a new and fascinating element for the BGB managers. A total of 188 accessions of rice have been assigned their DOI to be available online after the training held in Yogyakarta-Indonesia 2019 (https://glistest.planttreaty.org/glis/entity/list-own). This allows plant genetic resources users and providers to find more information on material transfer if necessary. Due to this new skill, two activities are underway of preparation in Burundi to increase a number of digitalized accessions: (1) Knowledge sharing session for the team in charge of managing BGB data and local collaborators; the experience sharing will focus on use of the software and the recording data on DOI and error handling; (2) An upgrading data on conserved germplasm, particularly missing information on PGRFA will be adjusted to facilitate future DOI registration.

Concluding Remark and Future Perspective

The Global Information sharing system on PGRFA offers easy access to information on PGRFA for research and breeding. Boosting the Big Data of PGRFA is a way to make data managers be aware on the use of digital object identifiers (DOIs) to find solution on problems related to the lack of information exchange on plant material. For Burundi, this approach can strengthens cooperation in term of biodiversity conservation. Registration with the DOI of data from the BGB is likely to improve the profile of the BGB PGRFA. The software tested to support the DOI application will facilitate the documentation and data sharing in Burundi. Promoting the use of DOIs suitable for identifying plant genetic material around the world will be performed in Burundi.

As the ITPGRFA facilitates access and benefit sharing through the establishment of information and genetic resources exchange scheme in a multilateral system, the BGB plan to: (i) set the germplasm data of all food crops in use for food or conserved according to required standard for registration with DOI. Preparation will be completed on each crop to overcome the challenge of lack of documentation on Burundi PGRFA and information sharing weakness on plant material kept in BGB or on farm; (ii) to set up a specific team in charge of data processing once approved by the scientific committee of ISABU. The team will be in charge of data processing and boosting for sharing information on GLIS.

REFERENCES

ISABU (2009) Regeneration and Safety Duplication of Regionally Prioritized Crop Collections. Annual Report of Bioversity International, Rome, Italy. ISBN 978-92-9043-830-4.

ITPGRFA (2017) Descriptors and guidelines of the Global Information System. ITPGRFA. http://www.fao.org/plant-treaty/areas-of-work/global-information-system/fr/.

Bashaasha, B. et al. (2010) Climate Change Vulnerability, Impact and Adaptation Strategies in Agriculture in the Eastern and Central Africa Region. Full Report. Makerere University, Michael Waithaka, Miriam Kyotalimye (ASARECA-PAAP). p. 14.

ISABU (2018) Rapport annuel sur la conservation et l'utilisation des ressources Phytogénétiques au Burundi. Organisation Mondiale Pour L'Alimentation et L'agriculture p. 71.

Bigirimana, et al. (2016) Résultats de la recherche action sur l'évaluation et la caractérisation des variétés d'igname, de niébé, de sorgho, du pois cajan et de taro dans la région naturelle de Moso pour faire face aux changements climatiques: Année agricole 2015-2016, p. 4-5.

GLIS (2019) Toolkit v.2.0.3 installation and operation manual. p 2 GIS: https://glistest.planttreaty.org/glis/entity/list-own.

PHILRICE GEMS DATABASE: ITS UPGRADE AND IMPACT

Malvin Duldulao, Marilyn Ferrer, Maria Cristina V. Newingham, Jonathan Niones, Xavier Greg Caguiat, Loida Perez, and Gabriel Romero

INTRODUCTION

Access to genetic resources and information held at genebanks are essential for current and future needs of rice improvement programs. Genebanks are responsible not only for the integrity of the samples they maintained but also with the data associated to the accessions (Richards and Volk 2015). The Philippine Rice Research Institute (PhilRice) Genebank serves as the national repository of 7.129 accessions of traditional landraces and other foreign-sourced rice genetic materials. It has enormous amount of data that was organized into a standard database format to enable strategic management for effective use and enhancement (Arlett et al. 2007). The Germplasm Management System (GEMS) was developed to facilitate access of information and to support complex processes in managing rice collections. It runs as a stand-alone database focused for internal use in the genebank operations. GEMS is also a central repository of all integrated data on passport, morpho-agronomic characterization, grain quality, pests, diseases, and abiotic stress reactions.

In order to optimize its use in rice conservation and management, it was upgraded into advanced version 'GEMS v2.0'. System architecture was redesigned, and new functions were added in synchronization with the current documentation needs of the genebank. This paper presents the significant impact of system upgrade from the perspective of genebank's operational efficiency

System Improvement

The overall architecture of GEMS was redesigned to cater system improvements: establishment of data validation rules and error-trapping mechanism, automation of procedures that covers the full range of activities in managing germplasm, implementation of barcode technology for seed inventories and distribution, and adaptation of Digital Object Identifier (DOI) for improved tracking of germplasm movement.

GEMS v2.0 is designed to document the two main operational procedures of the PhilRice Genebank: (i) genebank's management per se (acquisition, conservation, regeneration, inventory, and distribution), and (ii) germplasm management per se (passport, characterization, biotic, abiotic, and grain quality evaluation), which nested sub-operational procedures (Figure 1).

Figure 1. The sub-operational procedures in the genebank (a), and system architecture of GEMS v2.0 in stand-alone mode (b).

Genebank Operations

GEMS v2.0 serves as decision-support for genebank curators on daily operations to monitor seed inventory, determine regeneration priorities, generate germplasm status reports, and process germplasm requests. This transition ensures sufficiency of viable stocks and preservations of the germplasm's genetic integrity that would allow researchers and breeders to draw on a stable reference: a sample that would be, as much as possible, genetically unchanged since its storage (Peres 2016).

The inventory features of the GEMS allow proper documentation to monitor the quantity and viability of seeds to carry out regeneration and distribution. The system was designed to issue a warning as seed reaches critical level: below 50 grams in terms of amount and less than 85% in viability. It automatically highlights accessions to alert genebank curator to limit further distribution of such accessions until they are regenerated.

Inventory management adapted the barcoding system to eliminate labelling mistakes such as transposition of numbers which are common in hand-written labels. The most space-efficient linear symbol of Code 128 (subset C) was used as labelling paradigm. This machine-readable barcode label enables the system to acquire data via scanning seed packets thereby facilitating ease of updating seed stocks during seed distribution (Fig 2).

Seed and germplasm data requests processed primarily depending on the amount of seed stocks, viability status of the germplasm requested, and availability of information. This transaction requires the requestor to fill-up a "Request Form" and provide personal information, seeds/data to be requested, purpose and date of request, and acceptance to the Standard Material Transfer Agreement (SMTA) if applicable. All requests has assigned tracking number and documented in the 'Requests' tab of the system.

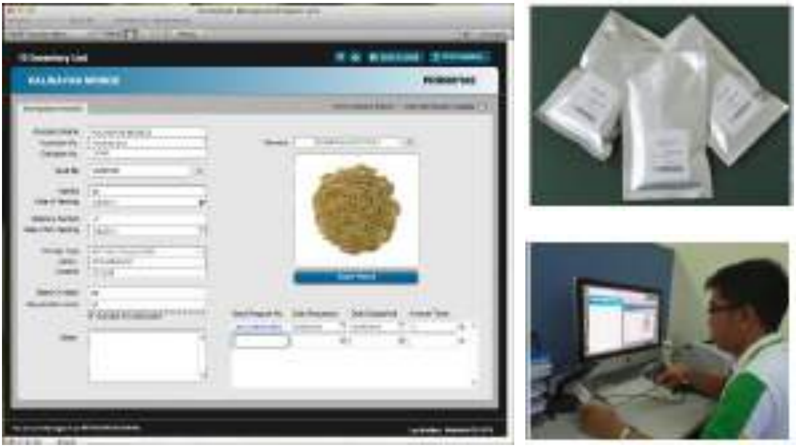


Figure 2. Scanning of seed packets with computer generated barcode labels.

Monitoring the use of PGR is one of the many responsibilities of genebank curators, but often this information is not reported or relayed back to the genebanks (Ruas 2017). GEMS v2.0 now provides capability to alert genebank curators to verify receipt of materials and record feedbacks after the expected time of delivery. Furthermore, the DOI-assigned germplasm will allow the use of material to be tracked, thus meeting the legal obligations of the SMTA and monitor the impact of genebank collections in utilization in research and breeding programs.

Germplasm Data

As part of data integrity checks, all data entry in the database is now executing validation rules for all parameters described for each germplasm collection. This mechanism ensures that dataset meets certain conditions before allowing it to commit the changes in the system. It automatically intercepts encoding errors and warns the user that invalid entry has been entered, thus, integrity and accuracy of rice germplasm data is maintained.

Passport data includes basic information of the collected material (*e.g* accession name, collecting number, collecting date, name of collector/donor, and place and date of collection, etc). To identify individual samples, a unique germplasm identification (GID) and collection number were assigned when passport data is encoded in the 'Registration' tab of the GEMS v2.0. Some other basic information consisting of 36 descriptor states will be encoded in the 'Collecting Data' tab (Fig 3). The system was also adjusted and complemented with DOI through registration to the Global Information System of the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) comprising of 1.016 accessions.

In addition to passport data, the database also contains morphological-data based from standard rice descriptors published by Alercia (2011). To facilitate ease of data entry and data viewing, the 'Characterization Data' tab of the GEMS v2.0 was organized into two groups (quantitative and qualitative data) separated by three views for vegetative, reproductive and post-harvest data (Fig 3). This consists of 58 parameters essential to facilitate preliminary selections of end-users for direct use or as parental in breeding programs. The system has the capability to automate the selection of materials to be included in the planting plan during Wet Season (WS) trials. Auto-selection of entries is based mainly on availability of morpho-agronomic characterization data, seeds stock and viability.

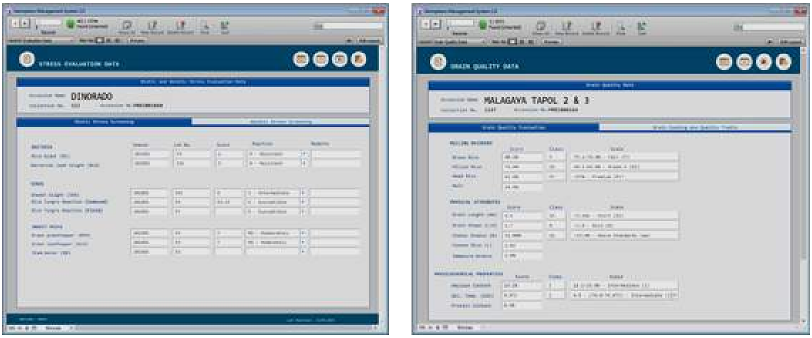
Table 1. List of parameters gathered at different rice growth stages.

Rice growth stage		Descriptor		
Vegetative				
	1	Basal leaf sheath color	7	Leaf blade pubescence
	2	Leaf sheath anthocyanin coloration	8	Auricle color
	3	Leaf blade anthocyanin coloration	9	Collar color
	4	Leaf blade distribution of anthocyanin	10	Ligule shape
	5	Leaf blade intensity of green color	11	Ligule color
	6	Leaf blade attitude (penultimate leaf)	12	Flag leaf (early)
Early reproductive				
	13	Culm habit (angle)	23	Awn distribution
	14	Culm kneeing ability	24	Awn color (early)
	15	Culm anthocyanin coloration on nodes	25	Panicle main axis attitude
	16	Culm underlying node color	26	Panicle branches attitude
	17	Culm internodes anthocyanin	27	Panicle secondary branching
	18	Culm underlying internodes color	28	Panicle exertion
	19	Flag leaf attitude (late)	29	Culm lodging resistance (culm strength)
	20	Stigma color	30	Heading days
	21	Lemma and palea color(early)	31	Days to maturity
	22	Lemma color of apiculus (early)		
Harvest and post harvest				
	32	Panicle shattering	44	Flag leaf width (cm)
	33	Panicle threshability	45	Culm length (cm)
	34	Awn color (late)	46	Culm number per plant
	35	Lemma and palea pubescence	47	Culm diameter at basal internode (mm)
	36	Lemma and palea color(late)	48	Panicle number per plant
	37	Apiculus color (late)	49	Panicle length of the main axis (cm)
	38	Sterile lemma color	50	Awn length (cm)
	39	Caryopsis: pericarp color	51	Sterile lemma length (mm)
	40	Endosperm type	52	Spikelet fertility
	41	Leaf length (cm)	53	Grain length (mm)
	38	Sterile lemma color	54	Grain width (mm)
	39	Caryopsis: pericarp color	55	Grain weight of 100 grains (g)
	40	Endosperm type	56	Caryopsis length (mm)
	41	Leaf length (cm)	57	Caryopsis width (mm)
	42	Leaf width of blade (cm)	58	Caryopsis shape (length/width)
	43	Flag leaf length (cm)		

Figure 3. Passport and morpho-agronomic data tab of GEMS v2.0.

Grain quality evaluation is one of the primary considerations for both consumers and breeders which will be encoded under 'Grain Quality Data' tab (Fig 4). This consist of data on milling recovery, physical attributes and physicochemical properties of the grain covering 14 parameters based from the grain quality classification on National Cooperative Test (NCT) Manual for Rice (2014).

Germplasm evaluations on abiotic and biotic stresses are significant in identifying resistant cultivars under unfavourable conditions. Data from the evaluation of germplasm against important biotic and abiotic stresses are to be encoded in the 'Stress Evaluation' tab (Fig 4).This was categorized into four categories namely: bacteria (rice blast, bacterial leaf blight), virus (sheath blight, rice tungro induced, rice tungro field reactions), insect pests (brown planthopper, green leafhopper, stemborer) and abiotic stresses (zinc, drought, salinity).



DINORADO

MALAGAYA TAPOL 2 & 3

Figure 4. Grain quality and stress evaluation data tab of GEMS v2.0.

The GEMS v2.0 was upgraded from FileMaker 4.0 (FM4.0) into a new cross-platform relational database-FileMaker Pro 12 Advanced and can be accessed internally upon request. The choice of using FileMaker software was prompted due to the previous data loss of in the Windows PC-based programs being plagued by software virus. As a commercial package, FileMaker Pro has extensive documentation (http://www.filemaker.com) and continually offers other features such as searching and report generation.

Corrections and changes will inevitably make in any database (McLaren et al. 2005). GEMS v2.0 has administration controls which allow assigning specific privileges to users. The system administrator has the 'admin privilege' and has the overall control to the system, while users have 'read-write privilege' that can modify but cannot delete records. Every change in a record is logged at 'System Logs' tab and can be undone if required. This feature protects the system from unauthorized modifications and possible security risks.

Concluding Remarks and Future Perspective

A new version of PhilRice Genebank in-house documentation system called 'GEMS v2.0' was launched in 2017. System upgrade was done wherein data validation was performed to ensure that system coding and source data migrated successfully. GEMS v2.0 with new system functions and user interface delivers significant impact in the genebank operational efficiency. Efforts are continuously being done to improve its data content and to ensure high quality germplasm data. To date, the web-based version of GEMS v2.0 is in alpha-testing and expected to access within the domain of PhilRice network by October 2019. Through this online interface, researchers and breeders can access information about the characteristics of rice varieties for rice varietal improvement programs.

REFERENCES

Alercia, A. (2011) Key Characterization and Evaluation Descriptors: Methodologies for the Assessment of 22 Crops. Bioversity International, Rome, Italy. ISBN 978-92-9043-874-8.

Correa-Victoria, F., Fiorella, E., Prado, G., & Aricapa, G. (2000) Population dynamics of the rice blast pathogen in a screening site in Colombia and characterization of resistance. 15: 214-220. Doi: 10.1007/978-94-015-9430-1_26.

McLaren, C.G., Bruskiewich, R.M., Portugal, A.M., & Cosico, A.B. (2005) The international rice information system. A platform for meta-analysis of rice crop data. Plant Physiology, 139(2):637-642. Doi:10.1104/pp.105.063438.

Peres, S. (2016) Saving the gene pool for the future: Seed banks as archives. Studies in Histology and Philosophy of Biological and Biomedical Sciences, 55: 96-104.

Richards, C.M. & Volk, G.M. (2010) New challenges for data management in genebanks. Acta Hortic, 859:333-336. Doi:10.17660/ActaHortic.2010.859.39.

Ruas, M., Guignon, V., Sempere, G., Sardos, J., Hueber, Y., Duvergey, H., Andrieu, A., Chase, R., Jenny, C., Hazekamp, T., Irish, B., Jelali, K., Adeka, J., Ayala-Silva, T., Chao, C.P., Daniells, J., Dowiya, B., Effa, E.B., Gueco, L., Herradura, L., Ibobondji, L., Kempenaers, E., Kilangi, J., Muhangi, S., NgoXuan, P., Paofa, J., Pavis, C., Thiemele, D., Tossou, C., Sandoval, J., Sutanto, A., Vangu, P.G., Yi, G., Van den houwe, I., Roux, N. & Rouard, M. (2017) MGIS: managing banana (Musa spp.) genetic resources information and high-throughput genotyping data. The Journal of Biological Database and Curation, bax046: 1-12. Doi: 10.1093/database/bax046.

MARDI AGROBIODIVERSITY INFORMATION SYSTEM (AGROBIS) TO SUPPORT GENETIC RESOURCES FOR FOOD AND AGRICULTURE (PGRFA) CONSERVATION IN MALAYSIA

Azuan Bin Amron, Site Noorzuraini Binti Abd Rahman,
Muhammad Izzat Farid Bin Musaddin,
Muhammad Luqman Hakim Bin Muhammad Fuad,
Mohamad Zulkifly Bin Zakaria,
Mohd Shukri Bin Mat Ali Ibrahim, Rosliza Binti Jajuli, and
Mohamad Zabawi Bin Abdul Ghani

INTRODUCTION

Malaysia is rich in biological diversity, harboring some 185000 species of fauna, more than 15000 species of flowering plants. Of about 1500 genera to be found are over 2500 tree species, 3000 species of orchids, 500 species of ferns, 60 species of grasses and bamboos, and many others (Abdullah et al. 2015). However, only a handful of the 15000 species has been utilized for food production. It has been reported that only about 300 species native to the country have been exploited and utilized (Ali et al. 2017).

Being the lead agency in spearheading the agricultural technology development in the country, MARDI has been identified as the national focal point for research and management of agrobiodiversity since the year 1996 (Abdul Shukor et al. 2013). Since then, Management and Utilization of Biological Resource Programme under the Strategic Resource Research Centre of MARDI have been given the task to undertake the national responsibility in the conservation, characterization, evaluation and utilization of crop genetic resources (Nordin et al. 2011). Genebank and Seed Centre were established after MARDI restructuring. The Management of Genetic Resources and Germplasm Conservation Program have been established to coordinate the management of genetic resources and germplasm at MARDI. Now, the management of all seed and germplasm collections at MARDI stations throughout the country is under the supervision of Agrobiodiversity & Environment Research Centre. The main objective within the programme is to mitigate the loss of agrobiological resources in the country through their judicious management and utilization in support of sustainable national food and agriculture development. In addition, the other functions of the programme are to gather and disseminate information on the nation's agrobiodiversity; to conduct research and development in areas relevant to management and sustainable utilization of agrobiological resources and to position MARDI as the national reference center for agrobiodiversity.

MARDI Initiatives on PGRFA Information Systems

Malaysia in implementing the Convention on Biological Diversity (CBD) that ratified in 1994 and currently, to support the Sustainable Development Goals (SDGs), always works on strategies in order to have adequate information on our biodiversity online. Agrobiodiversity research in MARDI, as a subset of biodiversity, generates great volumes of data. It comprises data from PGRFA, Animal genetic resources for food and agriculture (AnGR), and arthropods of importance to agriculture and microbial culture collection. In addition, this information is important for government at all levels, and also needed by users of the agricultural sectors, such as researchers, farmers, seed companies and education institutions. As a part of our National Biodiversity Strategies and Action Plans (NBSAPs), accessible information on this is online.

As for PGRFA particularly, in MARDI's genebank, researchers actively regenerate and multiply seeds, as well as screen PGRFA materials for different characteristics. These activities generate huge data. Before the year 2000, this data was not disseminated extensively and therefore the data was not accessible to breeders who are working for improving genetic resources. The information on characterization and evaluation work were normally kept only as hard copies, otherwise on the researcher's computer. MARDI realizes that was about time to supply this information and needs to develop its own information systems.

In 2002, MARDI has established Rice GeneBank Information System (RGBIS) and the system was developed for MARDI Rice Genebank operations. Subsequently, several individuals stand-alone were developed using Microsoft Access and Visual Basic 6 interface, such as *Nephelium lappaceum* (rambutan), *Mangifera odorata* (kuini), medicinal plants, arthropods and microbes. This database was a stand-alone database because of researchers involved wanted to protect unique characters of each agrobiodiversity group, especially for the plant group.

MARDI then developed a database information system known as MARDI Agrobiodiversity Information System (AgrobIS). The development of AgrobIS was taking into account that MARDI had several separate databases for specific projects or collections.

Apart from non-centered systems, these separate databases also make it difficult to access information. AgrobIS supports multiple collections of different datasets by using object-hierarchical design to enable inter-relational between each of agrobiodiversity components even with different data format such as a database, pictures and geographical data.

PGRFA data in AgrobIS do not follow the Multi-Crop Passport Descriptors (MCPD) because the data provider involved in the AgrobIS development is suggested for a system that enables them to develop their own descriptors based on the descriptors by IPGRI (Bioversity International). The central reference of the AgrobIS for PGRFA is the accession which is an anchor for the information belonging to the species. To give an impression about the scope of AgrobIS, the most important categories of Figure 1 is briefly described Tosiah (2013) as follows:

1. Passport data – contains basic information of the collection, such as, accession number, collector's name, collection date, cultivar name, and donor name
2. Environment and site – consist of data or information on sites where the accession is collected
3. Management – consists of information on how the collection being managed, duplicated area, preservation techniques and other related information
4. Characterization – data and information on characters of vegetative growth, leaf, inflorescence, fruits and seeds
5. Evaluation – data and information on pest status, fruiting times, biochemical contents, cytology characters and molecular markers.

Figure 1. PGRFA data categories in AgrobIS Interface for descriptors.

PGRFA cluster includes several components which are fruits, rice, traditional vegetables, herbs, medicinal plants, floriculture, palms and tubers groups. To date, there are 17.243 accessions were registered into AgrobIS. AgrobIS development has brought several benefits, especially in PGRFA conservation, mainly its encouraged utilization of conserved PGRFA by providing specific information of crop accession to the public that can be potentially used in crop improvement or as a new source for food and income. This also helps to support the key task of genebank which is to provide seeds to breeders, scientists and the public.

MARDI has developed The Seed Management Information System (SMIS) which is designed for MARDI Seed Genebank to manage the plant genetic resources stored in a seed genebank in the proper way. The system provides with a powerful data entry, flexible, and easy-to-use and linked to AgrobIS. It was developed in 2014 under the MEGA Project Fund. The objective of the project is to upgrade the system from "stands-alone" database, to the web-based system with easy access, integrated and comprehensive database system. This system contains information on varieties, quality attributes, seed regeneration, seed viability, and seed request status. This information is important to reduce the limitations and constraints in efficiently conserving the material in seed genebank and to allow easy access of the germplasm by researchers in MARDI or other universities and research institutes.

In the Eleventh Malaysia Plan (11 MP), MARDI has developed MARDI Rice Genebank Dashboard where allows breeders to search and retrieve information on specific accessions based on selected traits; and the information on rice germplasm were pulled from AgrobIS. The dashboard is able to provide information on rice germplasm focusing on high-yield, pest and diseases, and quality and specialty.

AgrobIS Data Available Globally

The Global Crop Diversity Trust (GCDT) has invited MARDI to adopt standards for data management and exchange through Genesys, which is the global online portal on PGRFA in the genebanks. In this project, MARDI's genebank has successfully developed an internal standard operating procedure (SOP) for preparing characterization and evaluation datasets for publication. Additionally, a very clear and relatable title to each of datasets, an explanation and description of the datasets, the location of where the research was conducted, the methodologies, the name list of researchers and supporting staffs involved in the creation of the dataset, and the precise descriptors used for the characterization or evaluation were created.

The rice germplasm data in AgrobIS were extracted and annotated in accordance with Multi-crop Passport Descriptors Version 2.1 (MCPD V.2.1) to build 22 rice metadata and datasets and published in Genesys catalogue. There 22 datasets that have been shared from AgrobIS into the Genesys portal: (1) Morphological characterization of rice accessions-off planting season 2012, (2) Morphological characterization of rice accessions-off planting season 2013, (3) Morphological characterization of rice accessions-main planting season 2013/2014, (4) Morphological characterization of rice accessions-off planting season 2014, (5) Morphological characterization of rice accessions-main planting season 2014/2015, (6) Morphological characterization of rice accessions-off planting season 2015, (7) Morphological characterization of rice accessions-main planting season 2015/2016, (8) Morphological characterization of rice accessions-off planting season 2016, (9) Morphological characterization of rice accessions-main planting season 2016/2017, (10) Evaluation of rice accessions for bacterial leaf blight (BLB) during the off planting season 2016, (11) Evaluation of rice accessions for bacterial leaf blight (BLB) during the off planting season 2015, (12) Evaluation of rice accessions for bacterial leaf blight (BLB) during the off planting season 2013, (13) Evaluation of rice accessions for bacterial leaf blight (BLB) during the off planting season 2011, (14) Evaluation of rice accessions for bacterial leaf blight (BLB) during the off planting season 2009, (15) Evaluation of rice accessions for bacterial leaf blight (BLB) during the main planting season 2015/2016, (16) Evaluation of rice accessions for bacterial leaf blight (BLB) during the main planting season 2014/2015, (17) Evaluation of rice accessions for bacterial leaf blight (BLB) during the main planting season 2011/2012, (18) Evaluation of rice accessions against Brown Plant Hopper (BPH), (19) Characterization of quality traits in rice, (20) Evaluation of rice accessions against Foliar Blast Disease (FBD), (21) Evaluation of rice accessions for bacterial leaf blight (BLB) during the main planting season 2013/2014, and (22)

Evaluation of rice accessions for bacterial leaf blight (BLB) during the main planting season 2009/2010.

Besides, MARDI has also involved in adoption and implementation of the Digital Object Identifier (DOI) which is a `permanent unique identifier created for all genetic resources registered in the Global Information System (GLIS). It is highly opaque and more robust than other permanent unique identifiers. Under the project *"W3B-PR-29-Indonesia: Multi-country 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"*, MARDI has registered the DOI for 727 rice accessions. The accessions that registered by DOI in MARDI rice germplasm includes MARDI released varieties, landraces, and introduced varieties from IRRI.

Concluding Remarks

In the era of Industrial 4.0 and Big Data, the AgrobIS will be continually upgraded with the current tools and systems. The awareness of the public, scientific community, especially the breeders and the agricultural scientists on the existence and importance of AgrobIS will be increased. Publishing AgrobIS data in the Genesys alongside with DOIs will help to contribute in the documentation of PGRFA and the way of how to exchange this information globally. The involvement and contribution of AgrobIS in both projects showed the commitment of MARDI for data sharing mechanism established by the Treaty.

REFERENCES

Abdul Shukor, A.R., Mohd Shukor, N. & Salma, I. (2013) Current and future strategies and practices in the conservation and sustainable utilisation of underutilised fruit genetic resources in Malaysia. Acta Horticulturae, 979:45-52.

Abdullah, M., Mamat, M.P., Yaacob, M.R., Radam, A. & Fui, L. H. (2015) Estimate the conservation value of biodiversity in national heritage site: a case of forest research institute Malaysia. Procedia Environmental Sciences, 30:180-185.

Ali, M.S.M., Ghazali, M.N.M., Khadijah, A., Azuan, A., Suryanti, B. & Maimun, T. (2017) Fungsi bank gen agromakanan kebangsaan (MygenebankTM), pusat ZZbank gen dan biji benih, MARDI dalam pemuliharaan sumber genetik tanaman untuk makanan dan pertanian. In Jajuli, R. & Ismail, N.A. (Eds.), Persidangan Agrobiodiversiti Kebangsaan 2016. Kuala Terengganu, Terengganu: MARDI.

Nordin, M.S., Ariffin, Z., Jajuli, R., Abdullah, W.D.W. & Denis, M.G. (2011) Conservation and Utilization of Plant Genetic Resources for Food and Agriculture in Malaysia. Serdang, Selangor: MOA.

Tosiah, S. (2013) AgrobIS: managing agrobiodiversity data and information in Malaysia. In Attaluri, S., Ajit M., & Gerard S. (Eds.), Agricultural Information and Knowledge for All: Success Stories on ICT/ICM in AR4D in Asia and the Pacific Region.

THE IMPLEMENTATION OF DIGITAL OBJECT IDENTIFIER ON GERMPLASMS AT THE GENEBANK OF CARDI

Phanna Phat, Vathany Thun, Lay Heng Seang,
Chan Sreypao Vat, Sophany Sakhan, and Ouk Makara

INTRODUCTION

Inherence from official concern voiced the need for conservation of plant genetic resources (PGR) in 1930, Food and Agriculture Organization (FAO) had promoted conservation, exchange and utilization of germplasms, and consequently set international guidelines in 1963. Rapid loss of diverse farmer landraces in 1970s forced into acceleration of PGR collection and establishment of genebanks which were expanded globally through the International Board for Plant Genetic Resources (IBPGR), formed by the Consultative Group on International Agricultural Research (CGIAR) (FAO 1997). Therefore, there are more than 1.700 genebanks around the world conserving over 7 million plant accessions (Singh 2012). In the context of Cambodia, the Cambodian Agricultural Research and Development Institute (CARDI) under the Ministry of Agriculture, Forestry and Fisheries of Cambodia has installed deep freezers for short-term conservation since 1989 for *ex situ* condition. Meanwhile, Genebank of CARDI stock approximately

7.000 accessions of rice germplasm and 1.600 accessions of wild rice and non-rice crops.

Genebank plays an important role in the conservation, availability and use of plant genetic diversity for crop improvement to meet the present and future demand for food production (FAO 2014). Genebanks make valuable resources utilizable for crop improvement through the acquisition, *ex situ* conservation and sharing of unique germplasm. Collecting plant genetic resources for food and agriculture (PGRFA) is important, however, the quality and management of the information associated with PGRFA is also as imperative as the material itself (https://www.genebanks.org/news-activities/news/dois/). Chauvet (1993) said "The efficiency of a genetic resources conservation network lies in the availability of materials and data." Some management aspects have been increasingly recognized as crucial to sustainable maintenance and optimal utilization of high-quality germplasm. Sharing accession information and making it publicly available for potential germplasm users is important to facilitate and support the use of the collection.

Every genebank created unique code for identifying each accession at theirown genebank. However, when contributing and sharing information of its PGRFA around the globe, there are problems with accession numbers which have the same identifying code but different materials. This is the global issue genebanks facing (https://www.genebanks.org/news-activities/news/dois/). Thus, the community agreed on the need of accurately and permanently identifying PGRFAs with global unique identifier. The Secretariat of ITPGRFA conducted a study on available technologies, a survey among over 200 experts worldwide, and a further validation study with 23 selected experts. Ultimately, the adoption of Digital Object Identifiers (DOIs) by the Scientific Advisory Committee (SAC) on the Global Information System. DOIs have now been chosen to provide a globally unique and permanent mechanism for identifying germplasm. The Secretariat of the International Treaty provides them free of charge for the entire PGRFA community participating in the Global Information System (FAO 2019). The Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD), in collaboration with the Secretariat of the ITPGRFA have lounged "The 2nd International Training on Software Testing to Support DOI Implementation" conducted on April 26th – May 3rd, 2019 with participants from 8 different countries including Cambodia. With the global information sharing of PGRFA assigned by DOIs, how Genebank of CARDI, Cambodia, respond to implementing DOI system in germplasm management and registration of Cambodian PGRFAs to Global Information System (GLIS).

Digital Object Identifiers

The DOI system has been developed and implemented in a range of publishing applications since 2000. DOIs were introduced by the International DOI Foundation and have largely been used for scholarly publications, datasets and commercial videos. It is widely used with over 190 million DOIs assigned (https://www.doi.org/faq.html). A DOI is a standardized alphanumeric string that is assigned by a registration agency and provides a persistent link to the location of information about the object on the Internet. Importantly, DOIs coexist with other identifiers, such as the ones already used by genebanks, allowing curators to keep their current systems in place. The Digital Object Identifiers (DOI) system provides identifiers which are persistent, unique, resolvable, and interoperable and capable of managing content on digital networks in automated and controlled ways (Paskin 2010). For these reasons, the ITPGRFA has made decision on adoption of Digital Object Identifiers (DOIs) by the Scientific

Advisory Committee (SAC) on the Global Information System for accurately and permanently identifying PGRFAs. DOIs have now been chosen to provide a globally unique and permanent mechanism for identifying PGRFAs (FAO 2019). Now DOIs hold great promise for tracking the use of germplasm (Paskin 2010).

Status of Genebank of CARDI

Isolation from rest of the world for 30 years started from late 60's to early 90's, Cambodia has never constructed a genebank (Sahai et al. 1992) until 1989 that the Cambodian Agricultural Research and Development Institute (CARDI) under the Ministry of Agriculture, Forestry and Fisheries of Cambodia has installed deep freezers for short-term conservation of rice germplasm with support from the Cambodian-IRRI-Australia Project (CIAP) (Javier et al. 1999). Cambodia would have a great irreplaceable loss of the native varieties of rice, if IRRI genebank did not receive the 55 viable rice varieties from Cambodia and later send a team to conduct the first collection, characterization, and conservation of 765 rice accessions at IRRI in 1972-1973 (Sahai et al. 1992). Through the CIAP project, CARDI has continued to collect, conserve, and characterize its traditional germplasms. Further,a total of 765 accession of first rice collection by IRRI and 2.842 accessions in four catalogues have been characterized, but there are remaining 3.400 samples which need to be characterized. Approximately 1.600 samples of wild rice and non-rice crops are collected and conserved in genebank of CARDI (CARDI 2018) (Table1).

Table 1. Wild and non-rice crops conserved in Genebank of CARDI in 2018.

No.	Common name	Scientific name	Conservation type	No of Samples
1	Wild rice	*Oryza rufipogon Griff.*	*Seed Bank*	295
2	barley	*Triticum aestivum*	*Seed Bank*	3
3	Maize	*Zea mays L.*	*Seed Bank*	58
4	Sorghum	*Sorghum bicolor*	*Seed Bank*	29
5	Okra	*Abelmoschus ficulneus*	*Seed Bank*	5
6	Sessame	*Sesamum indicum*	*Seed Bank*	13
7	Mungbean	*Vigna radiata*	*Seed Bank*	14
8	Soybean	*Glycine max*	*Seed Bank*	19
9	Groundnut	*Arachis hypogaea*	*Seed Bank*	16
10	Wild bean	*Vigna radiata*	*Seed Bank*	137
11	Watermelon	*Citrullus lanatus*	*Seed Bank*	43
12	Wax gourd	*Benincasa hispida*	*Seed Bank*	27
13	Pumpkin	*Cucurbita maxima*	*Seed Bank*	94
14	Cucumber	*Cucumis sativus*	*Seed Bank*	195
15	Sponge gourd	*Luffa acutangula*	*Seed Bank*	14
16	Bottle gourd	*Lagenaria siceraria*	*Seed Bank*	3
17	Bitter melon	*Momordica charantia*	*Seed Bank*	3
18	Wing bean	*Psophocarpus tetragonolobus*	*Seed Bank*	3
19	Tomato	*Solanum lycopersicum*	*Seed Bank*	37
20	Chili	*Capsicum annuum*	*Seed Bank*	180
21	Eggplant	*Solanum melongena*	*Seed Bank*	49
22	Mango	*Mangifera indica. L.*	*Field*	26
23	Fruit tree	Fruit tree	*Field*	30
24	Papaya	*Carica papaya*	*Seed Bank*	1
25	Cassava	*Manihot esculenta*	*Field*	28
26	Sweet potato	*Ipomoea batatas*	*Field*	36
27	Yam	*Oxalis tuberosa*	*Field*	1
28	Lesser yam	*Dioscorea esculenta*	*Field*	1
29	Potato	*Solanum tuberosum*	*In vitro*	4
30	Taro	*Colocasia esculenta*	*Field*	7
31	Banana	*Musa spp*	*Field + in vitro*	153
32	Sugarcane	*Saccharum officinarum. L.*	*Field*	34
33	Wild sugarcane	*Saccharum spontaneum*	*Seed Bank*	3
34	Cotton	*Gossypium hirsutum*	*Seed Bank*	1
35	Skouy	N/A	*Seed Bank*	1
36	Lotus	*Nelumbo nucifera*	*Seed Bank*	5
37	Ornamental plants	Ornamental plants	*In vitro*	5
	Total			1.573

Present Aspect of DOI Implementation in the GeneBank of CARDI

The 2nd International Training on Software Testing to Support DOI Implementation was conducted in Indonesia in order to introduce the DOI system, to transfer technical knowledge on DOI registration by using DOI Toolkit for managing rice genetic resources to genebank managers. Through the training, Genebank of CARDI has registered 10 accessions of rice germplasm on GLIS (Table 2). The DOI availability of Cambodian PGRFA in GLIS indexed in the ITPGRFA database will allow the genebank to preserve its database and facilitate international interest on exchange and utilization of Cambodian PGRFAs. The DOIs uniquely assigned to PGRFAs will provide breeders with a seamless way to find information and enable breeders and researchers to easily cite PGRFAs in scientific publication and documentation. This is useful for Cambodian breeders who are using our germplasms in breeding purpose and facing difficulty in citation and providing history of breeding materials. The remaining of crops germplasms are planned to be assigned their DOIs on GLIS.

Table 2. DOIs assignation of 10 accessions of Cambodian rice germplasms registering in the GLIS.

DOI	WEIWS code	Local ID	Creation method	Taxonomy
10.18730/SBCGX	KHM010	3458	Acquisition	*Oryza sativa* L.
10.18730/SBCHY	KHM010	3459	Acquisition	*Oryza sativa* L.
10.18730/SBCJZ	KHM010	3460	Acquisition	*Oryza sativa* L.
10.18730/SBCK*	KHM010	3461	Acquisition	*Oryza sativa* L.
10.18730/SBCM~	KHM010	3462	Acquisition	*Oryza sativa* L.
10.18730/SBCN$	KHM010	3463	Acquisition	*Oryza sativa* L.
10.18730/SBCP=	KHM010	3464	Acquisition	*Oryza sativa* L.
10.18730/SBCQU	KHM010	3465	Acquisition	*Oryza sativa* L.
10.18730/SBCR0	KHM010	3466	Acquisition	*Oryza sativa* L.
10.18730/SBCS1	KHM010	3467	Acquisition	*Oryza sativa* L.

Concluding Remarks

DOIs are the key to standardizing the identification of germplasm, facilitating information sharing and ensuring more effective use of the collection in genebank. PGRFAs with assigned DOIs will be easier to discover through digital means. The adoption of DOI system will not only help genebank of CARDI, but genebank around the globe to preserve its database online, to be able to track specific accessions to scientific publications and documents, and to share its PGRFAs with the world. Preserving PGRFA database on GLIS enable developing countries, where database managing system impossible to be installed and stop from data loss.

REFERENCES

Alercia, A., López, F.M., Sackville Hamilton, N.R. & Marsella, M. (2018) Digital Object Identifiers for food crops-Descriptors and guidelines of the global information system. Rome: FAO.

DOI (2019) Frequently Asked Questions about the DOI® System. Available from https://www.doi.org/faq.html.

FAO (1997) The state of the world's Plant Genetic Resources. Available from http://www.fao.org/3/a-w7324e.pdf [Accessed on June 8, 2019].

FAO (2014) Genebank Standards for Plant Genetic Resources for Food and Agriculture. Rev. ed. Rome.

CGIAR (2019) DOIs: A new era for germplasm labelling. Available from https://www.genebanks.org/news-activities/news/dois/ [Accessed on June 11, 2019].

Javier, L.E., Men, S., Pith, K.H., Khun, L.H., Say, P., Sin, S., Ouk, M., Hun, Y., Suy, S., Thun, V., Sidhu, G.S., Mishra D.P., Sahai, V.N., Chaudhary, R.C. & Ledesma, D.R. (1999) Rice germplasm catalog of Cambodia III. Cambodia-IRRI-Australia Project.

Paskin, N. (2010). Digital object identifier (DOI®) system. Encyclopedia of library and information sciences, 3:1586-1592. Doi:10.1081/E-ELIS3-120044418.

FAO (2019) Digital Object Identifiers (DOI). Available from: http://www.fao.org/plant-treaty/areas-of-work/global-information-system/doi/en/ [Accessed May 23, 2019].

Ouk, M., Sakhan, S., Seang, L.H., Thun, V., Kong, K., Orn, C., Nou, K., Vat, C.S.P., Tho, T., Houng, S., Ouch, S., Uong, P., Sen, S., Yon, S., Tat, C. & Sorn, T. (2017) Rice germplasm catalog of Cambodia IV. Phnom Penh, Cambodia: AFACI Project.

Sahai, V.N., Chaudhary, R.C. & Sin, S. (1992) Rice germplasm catalog of Cambodia I. Cambodia-IRRI-Australia Project. Phnom Penh, Cambodia.

Singh, A. K., Varaprasad, K. & Venkateswaran, K. (2012) Conservation costs of plant genetic resources for food and agriculture: seed genebanks. Agricultural Research, 1:223-239.

SIGNIFICANCE OF DIGITAL OBJECT IDENTIFIERS FOR PLANT GENETIC RESOURCES MANAGEMENT AND UTILIZATION

Asif Javaid

INTRODUCTION

Plant Genetic Resources for Food and Agriculture (PGRFA) are vital for modern plant breeding. They are essential in the efforts to increase agricultural productivity for food security. PGRFAs provide genetic traits required to deal with crop pests and diseases as well as dealing with changing climate conditions. They are also important for marginal environments in order to develop crop varieties that can adapt to difficult conditions. However, due to many reasons, plant genetic diversity is disappearing at alarming rates and 80-90 percent loss is reported in major crops over the past century (Andersen 2016).

Effective conservation and sustainable utilization of PGRFA are thus critically important for the food security of the world. *Ex situ* conservation of PGRFA began in mid-twentieth century when the importance of conservation of agricultural biodiversity was realized (Diez et al. 2018). Genebanks for PGRFA have the responsibility to collect, regenerate, conserve, characterize, evaluate, document and distribute germplasm (Tyagi and

Agrawal 2015). Genebanks also provide infrastructure for storage, a platform for sharing, and opportunity for better access and utilization of the germplasm. More than 1700 genebanks around the world stock over seven million plant accessions, of which only a small number are characterized and few are ever used for crop improvement. Since genebanks are expensive to establish and manage, there should be efficient genebank management system for the conservation and sustainable use of PGRFA (Singh et al. 2019). Utilization of PGRFA for research and breeding purposes requires thoroughly cataloged passport data and information about agronomic, biochemical and genomic traits of genebank accessions (Anglin et al. 2018). Modern phenotyping and genotyping technologies have enhanced the capacities of genebanks to perform their work but also raise questions as to how big data from plants should be managed and shared (Popova 2018).

Commercial use of PGRFA by corporate sectors is increasing, followed by demands of intellectual property rights. There are also demands for protection of the rights of farmers and local communities and ensuring a fair and equitable sharing of benefits arising from the use of these resources. The international community has responded to these developments with several regimes including the Convention on Biological Diversity (CBD), the Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) of the World Trade Organization and the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA). These regimes affect the management of PGRFA (Andersen 2016).

One important element of PGRFA management is accession number, which is a unique identifier assigned to each germplasm accession in a genebank. Since different genebanks have different systems of identifiers and data handling, many accessions in the genebanks are duplicates. Such information are often lost when the material is transferred from one holder to another. Different users of PGRFA, such as researchers and plant breeders, often follow different methods to assign identifiers to their materials. It results in a lack of standardization, which complicates the exchange of PGRFA data between researchers and plant breeders worldwide. The Global Information System of ITPGRFA aims to provide a standardized system of Digital Object Identifiers (DOIs) for plant genetic resources for food and agriculture (PGRFA) information around the world. It also facilitates easy access to information on PGRFA for research, training and plant breeding (Alercia 2018).

National Genebank of Pakistan

Pakistan is blessed with a vast diversity of plant genetic resources from different agro-climatic ecologies. Pakistan lies in the Central Asiatic Center of Diversity and in the vicinity of Near Eastern and Indian Centers of Diversity. Conservation of plant genetic resources at national level is required to supply the National Agricultural Research System (NARS) with germplasm. Germplasm collection was started in Pakistan during the 1970s. Bio-resources Conservation Institute (BCI) of Pakistan Agricultural Research Council (PARC) is established at National Agricultural Research Center (NARC), Islamabad, for exploration, collection, conservation, evaluation and distribution of PGRFA. More than a hundred plant collecting missions in Pakistan have collected indigenous germplasm from Pakistan for conservation and distribution. In addition, to meet the need and requirements of breeders and researchers, PGRFA were also acquired from abroad. National Genebank at BCI holds 38259 accessions of 400 plant species. These conserved genetic resources are distributed to breeders and researchers for utilization.

National Genebank of Pakistan distributes about 7000 accessions per year.

Germplasm Evaluation

Plant genetic resources without evaluation data do not serve the purpose of crop improvement. Identification of superior PGRFA saves the time and labor required for successful crop breeding. Germplasm evaluation also aids utilization and the planning for future collection missions. Evaluation Laboratory of BCI has characterized and evaluated more than half of the PGRFA conserved in the Genebank using standard protocols from Bioversity International. It identified genotypes for various traits of economic importance and enhanced the pace of crop development. Hidden treasures of germplasms tolerant to biotic and abiotic stresses have been uncovered. Markers for quality in wheat and disease resistance in peas have also been successfully utilized.

The Evaluation Laboratory has also identified elite crop germplasm lines with economically important characters like early maturity, yield components and resistance to biotic and abiotic stresses. National Genebank of Pakistan has made these elite germplasm lines available to education and research institutes for utilization in crop improvement.Biochemical evaluation of germplasm using SDS-PAGE and molecular evaluation using RAPD and SSR is also conducted to study the genetic diversity in germplasm of different crops.

Data Management and Integration

Genebank database maintains the passport data of germplasm accessions in National Genebank of Pakistan, while germplasm evaluation data is maintained by scientists working on different crops. The users of plant genetic resources conduct experiments in an effort to utilize germplasm to improve agricultural productivity. There is substantial variability in the quantity of experiments and the quality of data produced among different users. The potential uses of data from plant genetic resources are enormous for crop improvement. Future uses of agricultural trial databases will likely be driven by the increased links between genotype and phenotype to improve selection and use of germplasm. The combination of agronomic data from field trials with genomic data shows promise for developing next generation breeding. However, fragmentation, lack of organization and inaccessibility of agricultural trial data hinder their use and application for resolving problems in agriculture (Hyman et al. 2017).

Efforts are continuously being made for data integration. Several studies have suggested that Digital Object Identifiers (DOIs) can play an important role in data integration for management and use of plant genetic resources. Diez et al. (2018) discussed the strategies for improving Spanish Plant Genetic Resources Network, by identifying major shortcomings including the lack of efficient coordination in the distribution of species among genebanks, too many genebanks, existence of detected and undetected duplicates, insufficient rate of regeneration, insufficient phenotyping, genotyping and epiphenotyping, unsatisfactory rate of use by end users and insufficient funding. One of the suggested tasks is to develop user-friendly platforms to access germplasm documentation, including a unified system of descriptors and classification categories. Additionally, it was suggested that the use of digital object identifier (DOI) system would connect different data sources and increase the traceability of accession use.

Wambugu et al. (2018) reviewed the real and potential application of the current advances in genomic technologies in improving the utilization of plant genetic resources. The actual and potential application of these genomic approaches are in plant identification, phylogenetic analysis, analyzing the genetic value of germplasm, facilitating germplasm selection in genebanks, as well as instilling confidence in international germplasm exchange system. It was suggested that the genebanks can benefit from the current advances in genomic technologies through greater collaboration with the user community. Some of the linkages and collaborations that have been established include DivSeek and Global Information System (GLIS). The assignment of Digital Object Identifiers (DOIs) as permanent identifiers of PGRFA makes it easier to share PGRFA information, by easily and unambiguously referencing PGRFA samples across organizations.

The Wheat@URGI portal has been developed to provide the international community of researchers and breeders with access to the bread wheat reference genome sequence produced by the International Wheat Genome Sequencing Consortium. Genome browsers, BLAST, and InterMine tools have been established for in-depth exploration of the genome sequence along with additional linked datasets including physical maps, sequence variations, gene expression, and genetic and phenomic data from other international collaborative projects already stored in the GnpIS information system. The genetic resources stored in GnpIScoreDB also have the unambiguous identification of the accession using Digital Object Identifier (DOI) and a rich set of associated data (Alaux et al. 2018).

Genebank users should be in a position to identify and select germplasm from the global genepool based on a combination of passport, genotypic and phenotypic data. The Musa Germplasm

Information System (MGIS), the database for global *ex situ*-held banana genetic resources, has been developed to address those needs. Accession-based data has been enriched with publications, genotyping studies and associated genotyping datasets reporting on germplasm use. Banana germplasm identified by accession numbers will be complemented soon by Digital Object Identifiers (DOI) following the specifications of ITPGRFA (Ruas et al. 2017).

Concluding Remarks and Future Perspective

Efficient PGRFA database management system is necessary in genebanks. Genebanks use accession numbers for activities like collecting, regenerating, conserving, characterizing, evaluating, documenting and distributing germplasm. Once the accessions are distributed to users, their traceability becomes difficult because the user may use the same accession numbers or assign new accession numbers to the accession. DOI (Digital Object Identifier) is a unique identifier that identifies digital objects. Once DOI is assigned to a germplasm accession, the physical location of accessions may change, but the DOI assigned to that accession will never change. Thus DOI improves the traceability of accessions and their use. DOIs may also help resolve issues like intellectual property rights and access and benefit-sharing. Studies are being conducted on PGRFA for agronomic, physiological, biochemical and molecular characters by different research groups. Data and information generated during these studies should be integrated as in Wheat@URGI portal (Alaux et al. 2018) and MGIS (Ruas et al. 2017) using DOIs. It may help researchers and plant breeders to select PGRFA according to their goals.

The second International Training on Software Testing to Support DOI Implementation on Rice Germplasm was held in

Indonesia during April, 2019. During this training, GLIS Toolkit v.2.0.3 was used for assigning DOIs to rice germplasm accessions in National Genebank of Pakistan. The same software can be used to assign DOIs for germplasm accession from other crops. National Genebank of Pakistan is working on the development of new database management system for genetic resources. FAO Multi-crop Passport Descriptors (MCPD V.2.1) with DOIs as persistent unique identifiers are suggested to be used for development of new database management system (Alercia et al. 2015).

REFERENCES

Alaux, M., Rogers, J., Letellier, T., Flores, R., Alfama, F., Pommier, C., Mohellibi, N., Durand, S., Kimmel, E., Michotey, C. & Guerche, C. (2018) Linking the international wheat genome sequencing consortium bread wheat reference genome sequence to wheat genetic and phenomic data. Genome biology, 19(1):111.

Alercia, A., López, F.M., Sackville Hamilton, N.R. & Marsella, M. (2018) Digital Object Identifiers for food crops-Descriptors and guidelines of the Global Information System. Rome: FAO.

Andersen, R. (2016) Governing agrobiodiversity: plant genetics and developing countries. Routledge. p. 3-8.

Anglin, N.L., Amri, A., Kehel, Z. & Ellis, D. (2018) A case of need: linking traits to genebank accessions. Biopreservation and Biobanking, 6(5):337-349.

Alercia, A., Diulgheroff, S. & Mackay, M. (2015) FAO/Bioversity Multi-Crop Passport Descriptors V.2.1. Rome: Food and Agriculture Organization of the United Nations (FAO)-Bioversity International.

Díez, M.J., De la Rosa, L., Martín, I., Guasch, L., Cartea, M.E., Mallor, C., Casals, J., Simó, J., Rivera, A., Anastasio, G. & Prohens, J. (2018) Plant genebanks: present situation and proposals for their improvement. The case of the Spanish Network. Frontiers in plant science, 9:1794. Doi:10.3389/fpls.2018.01794.

Hyman, G., Espinosa, H., Camargo, P., Abreu, D., Devare, M., Arnaud, E., Porter, C., Mwanzia, L., Sonder, K. & Traore, S. (2017) Improving agricultural knowledge management: The AgTrials experience. F1000 Research, 6.

Popova, E. (2018) Special issue on agricultural genebanks. Biopreservation and Biobanking, 6(5):325-326.

Ruas, M., Guignon, V., Sempere, G., Sardos, J., Hueber, Y., Duvergey, H., Andrieu, A., Chase, R., Jenny, C., Hazekamp, T. & Irish, B. (2017) MGIS: managing banana (Musa spp.) genetic resources information and high-throughput genotyping data. Database, 2017:1-12. Doi:10.1093/database/bax046.

Singh, N., Wu, S., Raupp, W.J., Sehgal, S., Arora, S., Tiwari, V., Vikram, P., Singh, S., Chhuneja, P., Gill, B.S. & Poland, J. (2019) Efficient curation of genebanks using next generation sequencing reveals substantial duplication of germplasm accessions. Scientific reports, 9(1):650.

Tyagi, R.K. & Agrawal, A. (2015) Revised genebank standards for management of plant genetic resources. Indian J Agri Sci, 85:157-165.

Wambugu, P.W., Ndjiondjop, M.N. & Henry, R.J. (2018) Role of genomics in promoting the utilization of plant genetic resources in genebanks. Briefings in functional genomics, 17(3):198-206.

CHAPTER 5.
DIGITAL OBJECT IDENTIFIER FOR RICE GERMPLASM

Rice is one of the major food crops that contribute to food security and almost 20 percent of all the world's holdings of rice germplasm are included in the multilateral system of access and benefit sharing of the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA). Nevertheless, these accessions are maintained in collections distributed around the world and their discoverability and use is hindered by this fragmentation of location and database systems. Additionally, other important information associated with those rice accessions, which are highly relevant and required for plant breeding, is often disconnected from the collection database. The International Community has tried in the last decade to agree on a common approach to facilitate the establishment of an automated and meaningful links through the adoption of permanent unique identifiers.

Ambiguous germplasm identification and difficulties in tracing pedigree information also created constraints in rice improvement programs that involve exchanges of genetic resources. To address these constraints, a permanent unique identifier for each germplasm is needed. One such identifier is the Digital Object Identifier (DOI). The DOI is a permanent unique identifier that has specific advantages over other identifying systems, i.e. it is highly opaque and more robust than other permanent unique identifiers. Indonesia has applied the DOI to its rice germplasm collection based on the guidelines developed by the secretariat of the ITPGRFA. The DOI will keep track of the movement of Indonesian rice germplasms and can be used to further reveal the variations in the germplasms.

DIGITAL OBJECT IDENTIFIERS FOR INDONESIA RICE GERMPLASM

Muhamad Sabran, Nurul Hidayatun, and Hakim Kurniawan

INTRODUCTION

The mutual interdependence of different countries and regions in relation to plant genetic resources has been recognized by scientific communities (Palacios 1998). An in-depth analysis on interdependence on Plant Genetic Resources for Food and Agriculture (PGRFA) between countries demonstrated that national food supplies and production systems are highly interdependent worldwide in regard to plant genetic resources (Khoury et al. 2015). Countries strongly depend on crops whose genetic diversity largely originates from outside of their borders, both for their food supplies and their production systems. The global average of the degree of countries' dependence on crop genetic diversity originating from outside their borders is 68.7% across food supply variables, 69.3% across production variables, and 68.9% across all variables for all countries.

Acknowledging the inter-dependence among countries on plant genetic resources for food and agriculture, FAO member countries agreed to commit to an International Treaty on Plant

Genetic Resources for Food and Agriculture (International Treaty on Plant Genetic Resources for Food and Agriculture 2009). The Treaty established a Multilateral System (MLS) on access of 64 crops considered to be important for food security and climate change adaptation, and benefit sharing arising from their uses. In such multilateral system, contracting parties, international organizations and any natural and legal persons contribute to a virtual gene pool of the 64 crops. The transfer of genetic materials in the system are managed with a standard material transfer agreement (SMTA).

The SMTA requires that all available passport data and any other associated non-confidential descriptive information shall be made available in conjuction with the PGRFA provided. The recipient shall make available all non-confidential information that results from research and development carried out on the material to the multilateral system, through the information system specified in Article 17 of the treaty. However, rapidly growing material exchanges through SMTA had not yet included the sharing of associated information as required by Article 17 mentioned above.

Information System and Permanent Unique Identifier on PGRFA

There are many systems of inventory management and/or workflows of genetic resources available. One example is information systems in genebanks like GRIN GLOBAL and GENRES (https://www.genres.de/), that contains documentations on each accession and their origin, maintenance record of viability and health, genetic integrity of seeds or clones, etc. (Weise et al. 2017). Genesys (www.genesys-pgr.org), SINGER and

WIEWS (http://www.fao.org/wiews), and information systems for specific crops across genebanks such as ICIS and IRIS.

The global information system (GLIS) developed by the Treaty Secretariat aims to make such information more easily transferable, accessible, available for a long time, more accurately associated to the PRGFA at hand, and is accrued as the PGRFA is transferred, studied and improved. It is designed to become the "Google" for the PGRFA community and a gateway to existing systems for information on PGRFA and other resources of interest in the community. The global information system needs a permanent unique identifier for each accession so that they can be tractable and traceable.

A Permanent Unique Identifier (PUID), as a minimum definition, is a text string that unambiguously and permanently identifies a single object of interest (International Treaty on Plant Genetic Resources for Food and Agriculture 2015). The key characteristics of PUIDs are:

1. Uniqueness: unambiguously identify a specific object or intellectual asset
2. Permanence: the same object will be forever associated to the same identifier (always valid)
3. Opacity: nothing about the associated object should be inferable by the structure of the identifier
4. Actionability/resolvability: a defined procedure for name resolution exists to access the information associated to the object once the identifier is known
5. Discoverability: given details of an object, it is possible to retrieve its identifier

A Digital Object Identifier (DOI) (International DOI Foundation 2018) is a PUID with a number of specific advantages over other identifying systems, *i.e.* it is highly opaque and more robust than other PUIDs. The system is based on ISO and ANSI standards, thus ensuring a solid foundation and a controlled environment. It is managed by the International DOI foundation, widely used, accommodates existing identifiers, has a flexible and extensible metadata structure, and supports advanced features.

Guideline for the Use of DOI for PGRFA

In order to use DOI effectively, a consensus regarding minimum set of metadata descriptors for PGRFA has to be decided by the PGRFA communities. Through an electronic survey that involved 200 experts from 98 institutions in 62 countries, followed by several consultations, the community had agreed on a set of discriptors for assigning DOI to PGRFA. The data required for assignation of DOI (the descriptors) are grouped into core descriptors, which consist of mandatory descriptors and highly recommended descriptors, and additional descriptors (International Treaty on Plant Genetic Resources for Food and Agriculture 2017).

The values for mandatory descriptors must be inputted for every PGRFA material in GLIS. The holder is not given the option to enter “unknown”, “not applicable” or “other”, nor to leave the field empty. The PGRFA cannot be registered until valid values have been entered for all mandatory fields. These descriptor consist of organization, individual or legal entity conserving the PGRFA (MO1), PGRFA unique identifier within specific collection (e.g. genebank)(M02), the date at which the PGRFA was acquired (M03), the method by which the PGRFA was acquired (M04), and scientific or crop name (M05).

The method by which the PGRFA are obtained (M04) can be one of the following:

1. Acquisition: the acession was acquired from other sources, such as ex situ or in situ collection, local market, research collaborator, or any other sources outside the holding genebank, breeding or research program.
2. In-house copy: the PGRFA are created by subsampling or taking a harvest from another material that is under the holder's management, with the intent that the new PGRFA should be, as much as possible, a genetic copy of the parent, and the holder wishes to register a new DOI for the new PGRFA rather than use the DOI of the parent.
3. In-house variant: the PGRFA were created by subsampling or taking a harvest from another material that is under the holder's management. They are a subset of genetically variable (segregating and/or physical mixture) parents.
4. Novel distinct PGRFA: The PGRFA is a novel variety, breeding material, or research material that had been created from one or more parental samples under the holder's management, distinct from its parents, through a process that includes at least one innovative step such as crossing, mutation, or genetic modification.
5. Observation-Natural: the PGRFA material was in situ or on farm and it appeared on the holder's land without any intervention.
6. Inherited: The PGRFA were inherited from the holder's predecessor, without knowing how it was acquired or created.

There are also some highly recommended descriptors, although they are not mandatory, for the assignation of DOI such as : link to associated information, i.e. one or more URLs where further information about the PGRFA can be obtained (RO1), DOI(s) of progenitor(s) (RO2), biological status (RO3), additional taxonomic category, i.e., finer level of classification from the scientific names given in M05 (R04), names, i.e., one or more registered names or other designation such as a landrace, traditional variety or modern cultivar, or some other name or designation used to identify a breeder's selection or elite line or variety (R05), other identifiers (R06), MLS status, i.e. the status of the PGRFA with regard the multilateral system of ABS of the ITPGRFA (R07), Historical PGRFA, i.e., to indicate whether or not the material identified by the DOI currently exist (R08). In addition to those core descriptors, there are other descriptors that depend on the method by which the material are obtained as given in M04. Different values of M04 required different sets of additional descriptors. The DOI with the above descriptors for its assignation will facilitate access to the information about PGRFA.

An issue that was intensively discussed during the consultation is the definition of the "object" that needs to be identified with DOI for the purposes of the Global Information System. Defining the type of object to be identified requires a definition of what type or degree of difference (or change) is sufficient to require assignment of a new DOI. This needs a clear consensus for the term 'change'. More concretely, when a sample is transferred with SMTA from a provider to a recipient, should the provider's and recipient's samples be considered as the same object (and therefore share the same DOI) or to be different (and therefore have different DOIs)?

The provider's samples and recipient's samples from the same original material are separated following transfer, and for administrative, scientific, legal, and professional reasons must be treated as different objects. They are managed as separate entities under different administrative procedures, and unless the provider's sample was previously acquired with SMTA, they are held under different legal conditions. In addition, since genetic materials change through mutation, genetic drift, deliberate or unintentional selection, cross-contamination with pollen or seeds or plants of other varieties, or even simple mislabeling, the two samples will be genetically different; and they will diverge further with time. This is particularly true of landraces and crop wild relatives, which are typically genetically heterogeneous, so an accession or variety in a collection of genetic materials is often not a homogeneous sample of seeds. Documenting them as separate objects is essential to resolve questions that may arise over use or misuse, genetic identity, or attribution of data.

The main issue for discussion here is whether using the DOI for this purpose is in accord with the needs of the Global Information System. Choosing to assign the same DOI to the provider's and the recipient's samples implies a judgment that genetic differences between them are too small to warrant treating them as separate objects. Consequently, it would be necessary to agree to a threshold for the degree of genetic divergences between two samples where they can still be considered as the same object. Further agreements would need to be reached on what evidence would be required to determine whether they are different objects. Possible categories of evidence include (a) genetic characteristics such as DNA fingerprints, (b) phenotypic characteristics such as visible morphological traits, (c) records of germplasm management including sample sizes, and environments and processes used for germplasm maintenance and seed multiplication, or (d) the germplasm holder's subjective assessment. The issues described above have been resolved by allowing to assign DOI even to the in-house copy of the other accession as long as the method by which the PGRFA was acquired is reflected in the new DOI (in this case in-house copy, as outlined in descriptor M04).

International Rice Information System

The International Crop Information System (ICIS) (www.icis.cgiar.org), is a database system for the management and integration of global information on genetic resources and crop improvement. It was developed by CGIAR and its collaborative partners. The International Rice Information System (www.iris.irri.org) is the implementation of ICIS for rice (McLaren et al. 2005). The system consists of database design, controlled vocabularies, and software, which are applied to rice data. The IRIS contains the genealogy management system which documents the method of progeny creation, including generative (crossing/hybridization, induced mutation, and genetic modification), derivative (selection, separation of component from mixture), and maintenance (seed multiplication, sub-sampling) methods.

Linking DOI to the IRIS

The DOI assigned to rice germplasms needs to be linked with the IRIS (or ICIS in general) in addition to its registration in the Global Information System (GLIS). Providers submit a request to the GLIS through a web service by providing the local germplasm ID and passport data (according to the DOI metadata specification of GLIS). If no IRIS Germplasm Identifier Data (GID) is found in the provider's database, then a parent entry for the germplasm will be created in IRIS (in addition to the registration in GLIS) with minimal passport data and link it to the provider's database. The DOI server (GLIS) assigns a new DOI to the rice accession and this DOI is then added to: (1) the provider's own database and (2) the IRIS central database.

The provider submits a request to GLIS with DOI and recipient ID, and the system (GLIS) assigns a new DOI to the recipient’s sample and issues a service call to create a new entry for the recipient’s sample in IRIS with: GID, DOI, minimal passport data, and a link to the provider’s DOI. If the recipient’s database is enabled, a germplasm record with DOI and passport data is added to the recipient’s own database. If the material is not already registered, the provider can register it "on the fly" by providing the necessary information in the transfer request message. GLIS will first register the material, assign it a DOI and then proceed with the transfer as described above.

DOI for donor and recipient could perhaps be resolved by assigning separately one type of DOI where a type of object is defined as the original source material following the MCPD concept of COLLNUMB and the Darwin Core Occurrence concept (using the Darwin Core occurrence ID descriptor), and another type of DOI where its type of object is the actual germplasm material of the genebank-accession that is shipped in response to a seed request. The “COLLNUMB/occurrence ID-DOI” identifier name would be the same for both provider and recipient samples and correspond (in object-type definition) to the MCPD concept of COLLNUMB. The “germplasm-material-type-of-DOI” identifier name would be different for the provider and recipient samples and correspond (in object-type definition) to the core MCPD concept of ACCENUMB. This resembles the normal procedure of genebanks today where even a germplasm sample received from another genebank institute is assigned a new accession number (ACCENUMB) upon registration in the recipient genebank database. An obstacle for using this approach in GLIS is that the original source material (COLLNUMB) for genebank accessions is very often not known.

Another way to solve this is by exploring the different point of view between providers and recipients respectively. From the point of view of the provider, the provider's sample identifier (DOI) corresponds to the ACCENUMB identifier in the provider's own database, and the recipient sample identifier (DOI) will correspond to the "ACCENUMB-to-be" in the recipient's own information system. From the point of view of the recipient (and subsequent registration of the SMTA event in the recipient's own database) the provider's sample identifier corresponds to the MCPD DONORNUMB concept, and the recipient sample identifier (the material that is shipped) corresponds to the ACCENUMB concept with respect to registration in the recipient's own database. The GLIS SMTA component would thus need to take care to register both the provider's sample DOI and the recipient's sample DOI for each SMTA material transaction.

Documentation of passport data requires special attention, because it provides critical data or information for discriminating the germplasm materials. The higher the number of discriminants, the better its power to differentiate germplasm with DOIs. Therefore, the provider has to explore multiple potential traits to be used as reliable discriminant characters. In addition, as in Darwin Core, DOI is referenced for the present existence of germplasm materials in a genebank collection as well as the germplasm sample when it was first taken at the time of collection.

In the case of germplasm uses by contracting parties or other users, how can the DOI and other related data to the germplasm materials be tracked?

The Global information system that will be developed is expected to have the ability to trace the DOIs and related data to the germplasm materials exchanged by the MLS, and will compile these information for reporting to the MLS. Since the DOI will be centralized in GLIS (Global Information System), interested users and or any third parties will have the opportunity to further search and trace any related data to the germplasm material that is already shared in the MLS. It is important to note that accession number and DOI are the two different systems. Accession numbers are locally assigned and directly related to materials maintained in an *ex situ* germplasm collection, whereas DOI in this context is a globally unique identifier to be used to identify germplasm exchanged across multiple institutions.

DOI Applied to Indonesia Rice Germplasm

A successful rice breeding program requires genetic diversity to obtain improved variety with desired traits of economic importance. Germplasm characterization at morphological and molecular level will enable rice breeders to exploit genotypic diversity for conservation and further crop improvement (Thomson et al. 2009). Morphological variations can be used to classify rice into two major subspecies, japonica and indica. Typical grain characteristics and optimum environment between the two are distinctive. Another subspecies, the tropical japonica, has broad grains and thrives under tropical conditions and possesses specific morphology. Based on morphology and molecular analyses, scientists also try to sort rice into different cultivar groups, namely temperate japonica, aromatic, aus etc. (Garris et al. 2005). Morphologically, rice differs with respect to vegetative characters, leaf, culm, panicle, and seed properties during their growth stages, and these are important for guiding selection process in breeding.

Recent rice cultivars in Indonesia have more than 2000 ancestors and very complicated pedigree. IRRI cultivars contributed the largest proportion of the genetic background of Indonesian cultivars. Rice cultivars having the genetic background of IR64 or having kinship to IR64 were planted in 50.6% of the total rice cultivation area in Indonesia. Cluster analysis using coefficients of parentage divided Indonesian cultivars into 5 groups: Cisadane, Ciapus, IR8, Sintanur and Kalimas group.

Rice germplasms in Indonesia is managed by various institutions under the Indonesian Agency for Agricultural Research and Development (IAARD, Ministry of Agriculture), Indonesian Institute of Sciences (IIS), and Universities under the Ministry of Education and Culture. Rice germplasm collections in each institution are maintained because they are related to research and development activities. The largest rice germplasm collection is in IAARD with a total collection of 7,728 accessions. They are managed by the Indonesian Center for Rice Research (ICRR, 3,396 accessions) and Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD, 4,332 accessions). There is an additional ICABIOGRAD's collection of more than 8,000 accessions, which is deposited in IRRI genebank. The ICABIOGRAD genebank currently manages 4,332 accessions of rice, which consists of landraces (3,513 accessions), introduced varieties (357 accessions), wild rice (93 accessions) and others (369 accessions). ICABIOGRAD is the central genebank that holds a basic collection for rice germplasm. ICRR has a collection of 3,231 accessions, which consists of landraces (1,729 accessions), introduced varieties (1,287 accessions), wild rice (50 accessions), as well as elite varieties (165 accessions). Rice germplasm collection is also carried out by IIS and several universities.

However, the collection status in these institutions is still not completely inventoried.

Rice germplasm collection data in each institution is still managed using different database systems at the moment. There is no central database system that connects those institutions for accessing and sharing data. ICABIOGRAD has begun to coordinate rice germplasm data management in Indonesia. A prototype of website portal had been built since 2016 for the purpose of sharing agricultural genetic resources data. However, the coordination was still limited within the scope of the IAARD (Ministry of Agriculture).

Through the portal, it is possible to access collection and characterization data of agricultural genetic resources collected in ICABIOGRAD genebank. Until now, there have been a lot of access and exchanges of genetic resources material being carried out among the institutions within and outside the IAARD. Almost all of these material exchanges occured between domestic institutions.

Since 2017, a total of 1,366 accessions of rice and other crops collected in IAARD genebank have been registered to obtain the DOI. Most of those accessions were landraces, and they are maintained in ICABIOGRAD and some Assessment Institute for Agricultural Technology (AIAT). There are also minor accessions maintained in the gene bank of Agro-inovation Center, Gadjah Mada University (Table 1).

Table 1. Rice and non-rice accessions collected in IAARD gene bank which have been registered for the DOI.

Holding institution	Number of accessions	Year of DOI registration
ICABIOGRAD	849	2017
ICABIOGRAD	111	2018
AIAT of South Sumatera	55	2018
AIAT of Central Kalimantan	96	2018
ICABIOGRAD	70	2019
ICRR	14	2019
ILETRI	12	2019
ITFRI	10	2019
IRIAP	18	2019
ICSFRI	12	2019
ICERI	20	2019
AIC-GMU	30	2019
AIAT of Yogyakarta	55	2019
AIAT of West Nusa Tenggara	4	2019
AIAT of East Kalimantan	10	2019
Total	1,366	

PGRFA list

PGRFA (1-20 of 1366)

DOI ⇅	WIEWS code	Local ID	Date	Creation method	Taxonomy	Comm name
				Method... ▾		
10.18730/58TDE	IDN179	05020-03995	1970-09-24	Acquisition	Oryza sativa L.	Rice
10.18730/58TEF	IDN179	05020-04022	1970-10-05	Acquisition	Oryza sativa L.	Rice
10.18730/58TFG	IDN179	05020-04026	1970-10-05	Acquisition	Oryza sativa L.	Rice
10.18730/58TGH	IDN179	05020-04028	1970-10-05	Acquisition	Oryza sativa L.	Rice
10.18730/58THJ	IDN179	05020-04046	1970-10-13	Acquisition	Oryza sativa L.	Rice
10.18730/58TJK	IDN179	05020-04050	1970-10-13	Acquisition	Oryza sativa L.	Rice
10.18730/58TKM	IDN179	05020-04053	1970-10-13	Acquisition	Oryza sativa L.	Rice
10.18730/58TMN	IDN179	05020-04071	1970-10-13	Acquisition	Oryza sativa L.	Rice
10.18730/58TNP	IDN179	05020-04077	1970-10-13	Acquisition	Oryza sativa L.	Rice
10.18730/58TPQ	IDN179	05020-04079	1970-10-13	Acquisition	Oryza sativa L.	Rice
10.18730/58TQR	IDN179	05020-04082	1970-10-13	Acquisition	Oryza sativa L.	Rice
10.18730/58TRS	IDN179	05020-04084	1970-10-13	Acquisition	Oryza sativa L.	Rice

Figure 1. Recent status of DOI registration of the Indonesian PGRFAs accessions as shown in GLIS website.

PGRFA doi:10.18730/5BTDE

Citation: https://doi.org/10.18730/5BTDE

Organization/individual conserving the PGRFA: Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development
Jl. Tentara Pelajar 3A, Bogor 16111 West Java
Indonesia
WIEWS code: IDN179 [Details]
Easy-SMTA PID: 00AQ17

Local identifier: 05020-03995
Date: 1970-09-24
Creation method: Acquisition
Taxon: Oryza sativa L.
Common name: Rice

Biological status: Traditional cultivar/landrace
Names: Sengkumang
Other identifiers
MLS status: Annex 1 crop
Historical: No

Links to associated information

Keywords	URL
No links found	

©2019 International Treaty on Plant Genetic Resources for Food and Agriculture - FAO - Terms of use

Figure 2. Detail information for citation of a single accession of PGRFA.

As ICABIOGRAD-IAARD is a National Focal Point for managing agricultural genetic resources, therefore the DOI registration of domestic institutions was done under the coordination of ICABIOGRAD-IAARD. Technical assistances for DOI registration of agricultural germplasms collected in various institutions is still continuing even a collaborative project between IAARD and ITPGRFA-FAO has been ended in July 2019.

REFERENCES

Garris, A.J., Tai, T.H., Coburn, J., Kresovich, S. & McCouch, S. (2005) Genetic structure and diversity in Oryza sativa L. Genetics, 169 (3), 1631–1638. doi:10.1534/genetics.104.035642.

International DOI Foundation (2018) DOI Handbook.2018 [Online] Available from: https://www.doi.org/hb.html [Accessed: 16 August 2018].

International Treaty on Plant Genetic Resources for Food and Agriculture (2017) Data required for the assignation of Digital Object Identifiers in the global information system v.2.1.

International Treaty on Plant Genetic Resources for Food and Agriculture (2009) International treaty on plant genetic resources for food and agriculture. Rome, Food and Agriculture Organization of the United Nations.

International Treaty on Plant Genetic Resources for Food and Agriculture (2015) Technical options to facilitate the establishment of data links in the field of plant genetic resources for food and agriculture: permanent unique identifiers. San Diego.

Khoury, C.K., Achicanoy, H.A., Bjorkman, A.D., Navarro-racines, C., Guarino, L., Flores-palacios, X., Engels, J.M.M., Wiersema, J.H., Dempewolf, H., Ramírez-villegas, J., Castañeda-álvarez, N.P., Fowler, C., Jarvis, A., Rieseberg, L.H. & Struik, P.C. (2015) Estimation of co untries ′ interdependence in plant genetic resources provisioning national food supplies and production systems.

McLaren, C.G., Bruskiewich, R.M., Portugal, A.M. & Cosico, A.B. (2005) The international rice enformation system. A platform for meta-analysis of rice crop data. Plant Physiology, 139 (2), 637–642. doi:10.1104/pp.105.063438.

Palacios, X.F. (1998) Contribution to the estimation of countries ′ interdependence in the area of plant. Food and Agriculture Orgnization of the United Nations.

Thomson, M.J., Polato, N.R., Prasetiyono, J., Trijatmiko, K.R., Silitonga, T.S. & McCouch, S.R. (2009) Genetic diversity of isolated populations of Indonesian landraces of rice (Oryza sativa L.) collected in East Kalimantan on the island of Borneo. Rice, 2 (1), 80–92. doi:10.1007/s12284-009-9023-1.

Weise, S., Oppermann, M., Maggioni, L., Van Hintum, T. & Knupffer, H. (2017) EURISCO: the European search catalogue for plant genetic resources. Nucleic Acids Research, 45 (D1), D1003–D1008. doi:10.1093/nar/gkw755.

DIVERSITY OF DOI REGISTERED RICE LANDRACES FROM SOUTHEAST ASIA FOR GENETIC IMPROVEMENT

Dwinita Wikan Utami, Puji Lestari,
Rerenstradika Tizar Terryana, and Karden Mulya

INTRODUCTION

Rice is the most widely consumed as a staple food for a large part of the world's human population, especially in Southeast Asia. This plant species is the agricultural commodity with the third-highest worldwide production. Currently, the majority of all rice production comes from China, India, Indonesia, Bangladesh, Vietnam, Thailand, Myanmar, Pakistan, Philippines, Korea and Japan. Asian farmers including those in Southeast Asia still yield for 87% of the world's total rice production (Worldatlas 2017). Rice landrace is defined as a traditional variety with a high capacity to tolerate biotic and abiotic stresses, resulting in high yield stability and an intermediate yield level under a low input agricultural system. Landraces covering a domesticated, locally adapted (Jones et al. 2008) or traditional variety (Camacho et al. 2005) have developed over time, through adaptation to its natural and cultural environment of agriculture and pastoralism, and due to isolation from other populations (Sponenberg 2000).

Rice landraces possess wide diversity, which needs to be properly characterized for their use in genetic improvement. They have been preserved by indigenous farmers to the commercially bred cultivars developed during the green revolution. Harlan (1975a) was stated that landraces are balanced populations in equilibrium with both the environment and pathogens, and are genetically dynamic. Each landrace has particular properties or potential characteristics: early maturity, adaptation to particular soil types, resistance or tolerance to biotic and abiotic stresses, and increased grains number which potential for crop improvement through breeding approach (Das et al. 2013). For this reason, the management of the information of landraces is as important as the material itself associated with genetic resources for food and agriculture. The complicated problems arise when some landrace accessions were shared to different place. It becomes difficult to track relationships as these genetic resources are shared, duplicated and further used, causing limitation of the capacity to associate new information to the accessions attributable to generating it. Digital Object Identifiers (DOI) has now been chosen to provide a globally unique and permanent mechanism for identifying germplasm which is important to multilateral system. A DOI is a standardized alphanumeric string that is assigned by a registration agency and provides a persistent link to the location of information about the object on the Internet (unlike URLs, which can lose connectivity if they are not updated). Importantly, DOIs coexist with other identifiers, such as the ones already used by gene banks, allowing curators to keep their current systems in place (https://www.genebanks.org/news-activities/news/dois/).

This review described how the diversity of DOI registered Asian landraces, particularly those originating from Southeast Asia can be exploited and made readily available for use by breeders and the scientific communities. It highlighted DOI as a mean of research in genetic improvement to traceable the rice genetic resources for held in trust. An attempt was made to discuss the diversity of rice landraces in Southeast Asia on specific-trait characters and a special case of the diversity of shared rice landraces from multiple countries could be utilized through breeding for genetic improvement. Finally, ways of centralizing and unifying genetic resources and their potential associated traits for specific trait characters through the conservation of germplasm and the development of databases are proposed. It is clear that more efforts are necessary to exploit genetic diversity of rice landraces in Southeast Asia to achieve sustainable production in the future.

Genetic Diversity of DOI Registered Rice Landraces on Specific-Trait Characters

Rice landraces DOI registered diversity

The genetic variability found within landraces affords the possibility of genetic flexibility. Landraces have the potential to adapt to local field conditions, changing environments and farming practices (McCouch 2004). Moreover, the genetic diversity of traditional/landrace varieties is the most immediately useful and economically valuable component of rice biodiversity (Wood and Lenné 1997). To efficiently conserve, manage and use such germplasm resources, an understanding of structure, apportionment and dynamics of local landrace variation is required. Several studies have examined genetic variation and differentiation among rice landrace varieties (Fukuoka et al. 2006).

Germplasm identification using DOI has been globally adopted. DOI system for genetic resources, including rice landraces adopted by Scientific Advisory Committee (SAC) on the Global Information System (GLIS) has been applied in the management and the exchange of plant genetic resources for food and agriculture. This system has also been implemented on the total 85 accessions rice landraces from Southeast Asia countries, under 'Benefit-Sharing Fund of the International Treaty on Plant Genetic Resources for Food and Agriculture' Project funded by European Union through multi country project of FAO. All these accessions were registered into DOI system as showed on: http://biogen.litbang.pertanian.go.id/codevelopment/list-of-the-indonesian-local-cultivars-of-rice and in the multilateral system (https://ssl.fao.org/glis/). They were originated from: Lao PDR (17 accessions), Philippines (17 accessions), Malaysia (19 accessions) and Indonesia (20 accessions) and 3 check varieties from the four countries (total 12 varieties) which have been characterizedand included in the 3000 Rice Genome Project in International Rice Research Institute (IRRI) database (http://iric.irri.org/resources/3000-genomes-project).

Observation of all accessions from each country was diverse based on morphological characters. Based on principle component analysis, several characters contributed the most on total diversity (Figure 1). The characters and their contribution (%) are plant height (72.5%), length of ligule leaf (79.4%), color of leaf sheath (92.7%), color of collar leaf (85.7%) and color of auricle leaf (87.9%).

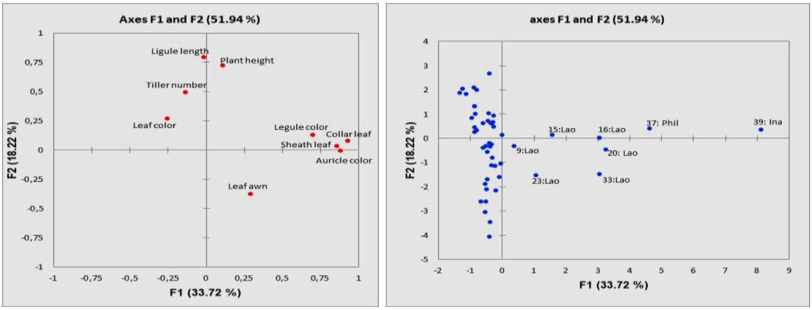


Figure 1. A. A distribution pattern of the morpho-agronomical characters contributed to the diversity on all accessions rice landraces from four Southeast Asian countries.; B. Relative positions of the accessions of rice landraces from four Southeast Asian countries: Mal=Malaysia; Ind: Indonesia; Phil=Philippine and Lao PDR.

Based on Principal Coordinate Analysis (PCoA) plotting, in general, landraces from the Philippines were clustered in same quadrant with landraces from Lao PDR, indicating relatively close genetic distance. While Malaysia landraces were clustered in the same quadrant with landraces from Indonesia. The accessions plotted in the same or closely coordinates indicated in low diversity (Mahbub et al. 2016). Therefore the PCoA plotting was applicable to analyse the genetic diversity among the accessions (Shankar et al. 2009). The closely related landraces from the two countries apparently reflect their ecological similarities (Burns and Strauss 2011).

Specific-traits characterization related with the breeding targets

The integration of physical and biological characters is the ultimate goal for characterization, as this would lead to an improved assessment of where particular types of environments are present (McLaren and Wade 2000). Morphological characterization generates important information that serves as the basis for selecting the accessions, either for direct use by farmers, use as parents in breeding programs or for varied scientific investigations, including gene discovery (Engels and Mba 2014).

Since physiological understanding of patterns of genotype adaptation is required for the identification of useful traits conferring an adaptive advantage in particular conditions, using the representative reference lines could provide the first step in the characterization process (McLaren and Wade 2000). The agro-climatic classification for rice and rice-based cropping systems has been widely adopted. This is based on the length of the rice-growing season, months in which surface flooding and disease incidence can be maintained (Oldeman and Frere 1982).

Southeast Asia region's hot and humid climate during the long and heavy monsoon season provide the most favourable agro-ecological environment for rice cultivation as well as diseases development. Cianjur district in West Java, Indonesia, is one specific lowland environment which is appropriate for rice landraces characterization due to the endemic status in all season for Bacterial Leaf Blight (BLB) disease. Therefore, this environment is appropriate for characterization of rice landrace accessions from Southeast Asia on the specific trait of BLB resistance (Figure 2).

Figure 2. Characterization of rice landraces from Southeast Asia in BLB's endemic lowland in Cianjur, West Java, Indonesia.

BLB is found to be one of the most destructive diseases of rice in many Southeast Asia countries. The disease starts with water soaked lesion on the tip of the leaves, and the lesion increases in length downwards and turns into yellow to straw coloured stripes with wavy margins. Lesions may be developed at one or both edges of the leaves or along the mid rib. In humid areas, on the surface of the young lesions, yellowish, opaque and turbid drops of bacterial ooze may be observed in the early morning (Mew and Gonzales 2002). This disease becomes a serious problem because many high yielding varieties, which are managed with high nitrogen levels and close spacing, have inadequate resistance to the pathogen.

Characterization for specific trait in BLB's endemic area of rice landraces originated from Lao PDR, Philippines, Malaysia and Indonesia showed diverse performance. Based on standard scoring produced by IRRI (2002) for BLB's resistance (score 1 to 9), around 40% from the total accessions were resistance (score 1 and score 3) to natural inoculum in Cianjur's environment (Figure 3A). It seems the diversity was generated by spatial factors because they are originated from different countries with different environment. Frankel and Soule' (1981) reported that one of the former on genetic diversity of landraces was attributable by diverse spaces and reproductive isolation. The rice landraces accessions characterized were separated from different places and constrained on their genetic transfer between each accession.

Among the resistance cluster, there were also diversity between origin of country with score 3 resistance accession being dominant (Figure 3B). Rice landrace accessions from Lao PDR which have the major resistance (score1) were the most susceptible comparing those from other countries. Notable, the genetic diversity of landraces within populations was generated by heterogeneity in time associated with both short term variations between seasons and by longer-term climate, biological and socio-economic changes (Frankel and Soule' 1981).

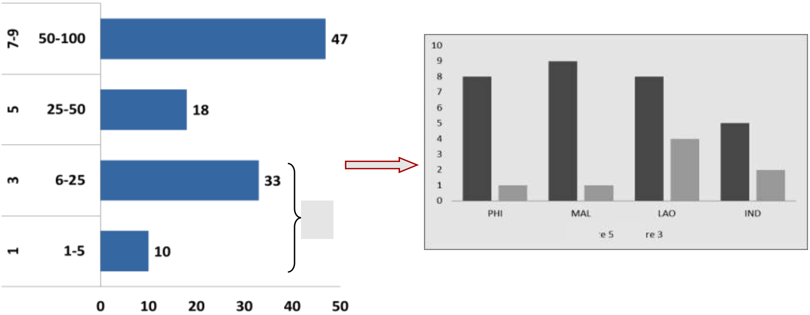


Figure 3. A. Distribution of BLB's resistance responds on rice landraces from Southeast Asia. B. The resistance group (s-1:score1 and s-3: score 3) in each country.

Finding Allelic Variation for Targets Traits Among Landraces

Conservation of all gene pools is a high priority for sustaining food security and coping with current and future climate change effects. Not only must landraces be conserved, but so should local varieties that have been replaced by new and more productive ones. Older varieties, due to the emphasis on landraces and more exotic materials, must not be forgotten, as well as other breeding materials, need to be conserved as a source of genetic diversity. Despite the enormous efforts made by national and international programs to conserve landrace diversities, eventually the conservation of germplasm and characterization of key traits will provide specific information to breeders that will promote the use of genetic resources by the scientific community.

Allele mining, being a tool at identifying allelic variation of relevant traits, is important to identify genes of known function and basic DNA sequence, therefore, genetic resources collections may be screened for allelic variation (Bhullar et al. 2010) using different molecular technologies. Isolation of important alleles from landraces and other genetic resources has been referred to as a source of functional allelic diversity and differential allelic frequency in comparison to modern breeding lines (material selected through breeding but not yet a variety) and varieties. This information will be highly valuable for enriching the genetic diversity within breeding programs. Discovery of new markers associated with key traits through genome-wide association studies (GWAS) will greatly benefit the scientific community, particularly after validation of important markers associated with complex traits (Lopes et al. 2015; Sukumaran et al. 2015).

Allele mining for BLB resistance on rice landraces was accomplished by specific markers for dominant major gene of

BLB resistance, Xa7, based on previous genomic map: 27.7-28.0 Kb on chromosome 6 (Chen et al. 2008; Proter et al. 2003; Utami et al. 2017). These specific markers could be applied to screen the rice accessions for allele mining purpose (Figure 4).

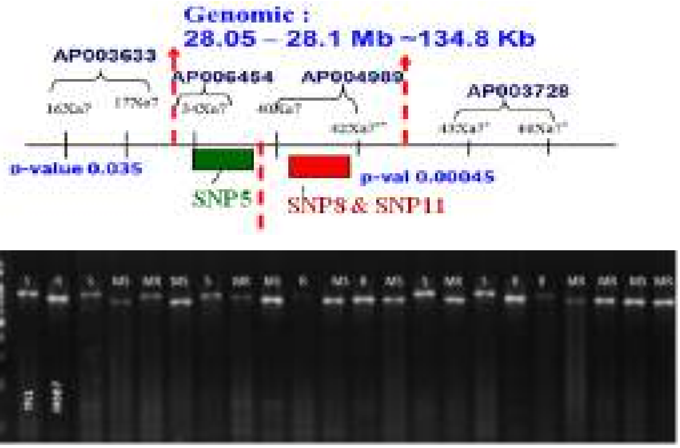


Figure 4. Allele mining of Xa7 gene by using the specific designed molecular markers based on fine map of the gene.

DOI Registered-Genetic Diversity and Allelic Variation into Gene Bank Database

Genetic diversity is the main source of variability in any crop improvement program. It serves as a reservoir for identifying superior alleles controlling key agronomic and quality traits through allele mining/association mapping. The assessment and conservation of diversity landraces have to be based on allelic variation at one or more loci. It will link to characterization activity to find the unique combinations of alleles which result in unique or distinct phenotypes. The discovery of abiotic stress at tolerant alleles in landraces of rice clearly shows the importance of conserving and exploring landrace germplasm as a mean to identify genomic based to find beneficial alleles for enhancing adaptation and productivity in stress-prone environments (Dwivedi et al. 2016).

The characteristics of landraces in relation to the magnitude of allelic and genetic diversity in contrast to cultivars are considered to be significantly more genetically diverse (Fowler and Mooney 1990). Thus, a landrace is a 'highly variable population in appearance' (Harlan 1975b), 'highly diverse populations and mixtures of genotypes' (Hoyt and Brown 1992), 'genetically heterogeneous' (Villa et al. 2006), 'not genetically uniform and containing high levels of diversity' (FAO 1998), 'local diverse crop varieties' (Brush 1999), 'heterogeneous crop populations' (Hoyt and Brown 1992) and 'materials with variable levels of heterogeneity'. The diversity of landraces was influenced by allelic variation which as a key priority traits in breeding programs.

The ability of accession landraces which have DOI registered is potential to undergo rapid adaptation or utilization for genetic improvement. As consequence, the utilization of rice landraces for genetic improvement could be traceable the rice genetic resources for holding it in trust and confidence. The next important thing is the DOI's label was coexisting with other identifier, such as in gene bank management.

Concluding Remarks and Future Perspective

Loss of genetic diversity has been recognized as a genetic bottleneck imposed on crop plants during domestication and through modern plant-breeding practices. Allelic variation of genes originally found in the wild but gradually lost through domestication and breeding has been recovered only by going back to landraces. Landraces have potentially new allelic variation that should be exploited in rice breeding.

Several strategies to retain diversity found in rice landraces from Southeast Asia are available and must be implemented. The first strategy is measuring diversity to establish core collections where genetic diversity is maximized with minimum repetition and in intermediate generations of a breeding programme to conserve genetic variability for selection in later generations. The second is addressing the allelic variation for key traits in breeding programmes where the importance of keeping diversity has been well established. Allele mining for traits and alleles of interest (finding 'new' allelic variation for known functional genes among landraces and linkages or pleiotropic between new alleles for known functional genes and grain yield) will be highly valuable for enriching the genetic diversity within breeding programmes. Phenotypic variation and related genetic association for specific traits through large-scale and precision phenotyping can be coupled with GWAS for the identification of new markers. The third strategy is using the DOI's label, utilization of rice accession landraces which could be traceable in trust and worthwhile for gene bank management. Finally, the unified development of databases and promotion of data sharing among physiologists, pathologists, scientists, national programmes and breeders through linkages with gene banks will greatly benefit on rice improvement for adaptation to climate change worldwide.

Acknowledgment

Authors thank to European Union to sponsor this BSF-ITPGRFA/FAO project (W3B-PR-08-Indonesia) which support data as the main part of this review article. Authors also give highly appreciation to technicians for their assistance to collect data from field assay.

REFERENCES

Bhullar, N.K. et al. (2010) Wheat gene bank accessions as a source of new alleles of the powdery mildew resistance gene Pm3: A large scale allele mining project. BMC Plant Biology, 10(88):1-13. Doi:10.1186/1471-2229-10-88.

Brush, S.B. (1999) The issues of in situ conservation of crop genetic resources. In: Brush,S. (ed.) *Genes in the Field*. Rome: International Plant Genetic Resources Institute. pp 3-26.

Burns, J.H. & Strauss, S.Y. (2011) More closely related species are more ecologically similar in an experimental test. *Proceedings of the National Academy of Sciences,* 108(13):5302-5307. Doi:10.1073/pnas.1013003108.

Camacho, V. et al. (2005) Defining and identifying crop landraces. *Plant Genetic Resources*, 3(3):373-384. Doi:10.1079/PGR200591.

Chen, S. et al. (2008) High-resolution mapping and gene prediction of *Xanthomonas Oryzae* pv. *Oryzae* resistance gene Xa7. *Molecular Breeding*, 22(3):433-441. Doi:10.1007/s11032-008-9187-1.

Das, B. et al. (2013) Genetic diversity and population structure of rice landraces from Eastern and North Eastern States of India. BMC Genetics, 14:1-14. Doi:10.1186/1471-2156-14-71.

Dwivedi, S.L. et al. (2016) Landrace germplasm for improving yield and abiotic stress adaptation. *Trends in Plant Science,* 21(1):31-42. Doi:10.1016/j.tplants.2015.10.012.

Engels, J. & Mba, C. (2014) Review of AFRICA rice Genebank, Cotonou, Benin Republic. Available from: https://www.genebanks.org/wp-content/uploads/2017/01/AfricaRice-genebank-review-2014.pdf.

FAO (1998) *The state of the world's genetic resources for food and agriculture*. Rome, Italy.

Fowler, C. & Mooney, P. (1990) *Shattering: Food, Politics and the Loss of Genetic Diversity*. Tucson: University Arizona Press.

Frankel, O.H. & Soule', M.E. (1981) *Conservation and Evolution*. Cambridge: Cambridge University Press.

Fukuoka, S. et al. (2006) Diversity in phenotypic profiles in landrace populations of Vietnamese rice: a case study of agronomic characters for conserving crop genetic diversity on farm. *Genetic Resources and Crop Evolution*, 53(4):753-761. Available from: https://link.springer.com/article/10.1007/s10722-004-4635-1.

Harlan, J.R. (1975a) *Crops and man*. Madison, Wisconsin, USA, American Society of Agronomy and Crops Science Society of America. Available from: https://www.cabdirect.org/cabdirect/abstract/19931634635.

Harlan, J.R. (1975b) Our vanishing genetic resources. Science, 188(4188):617-621. Doi:1.1126/science.188.4188.617.

Hoyt, E. & Brown, S. (1992) *Conserving the wild relatives of crops*. Rome: International Board for Plant Genetic Resources.

International Rice Research Institute (IRRI) (2002) *Standard evaluation system for rice* (SES). International Rice Research Institute.

Jones, H.L. et al. (2008) Approaches and constraints of using existing landrace material to understand agricultural spread in prehistory. *Plant Genetic Resources,* 6(6):98-112. Doi:10.1017/S1479262108993138.

Mahbub, M.M. et al. (2016) Morphophysiological variation in soybean (*Glycine max* (L.) Merrill). *American-Eurasian J. Agric. & Environ. Sci,* 16(2):234-238. Doi:10.5829/ idosi.aejaes.2016.16.2.12687.

McCouch, S. (2004) Diversifying selection in plant breeding. *PLoS Biology*, 2(10):1507-1512. Doi:10.1371/journal.pbio.0020347.

McLaren, C.G. & Wade, L.J. (2000) Using reference lines to classify multi-environment trials to the target population of environments and their potential role in environmental characterization. In: Tuong,T.P. et al. (eds.) *Characterizing and Understanding Rainfed Environments,* 5-9 Dec 1999. Bali, Indonesia. Los Banos (Philippines): International Rice Research Institute. pp. 131-144.

Mew, T.W. & Gonzales, P. (2002) *Handbook of Rice Seedborne Fungi*. Los Banos, Philippines: International Rice Research Institute.

Oldeman, L.R. & Frere, M. (1982) A study of the agroclimatology of the humid tropics of South-East Asia. *WMO Interagency Project on Agroclimatology*.

Proter, B.W. et al. (2003) Development and mapping of markers linked to the rice bacterial blight resistance gene Xa7. *Crop Science*, 43(4):1484-1492. Doi:10.2135/cropsci2003.1484.

Shankar, R. et al. (2009) Diversity analysis of bitter gourd (Momordica charantia L.) germplasm from tribal belts of India. *The Asian and Australian Journal of Plant Science and Biotechnology*, 3(1):21-25.

Sponenberg, D.P. (2000) Genetic resources and their conservation. In: Ann,T. & Ruvinsky,A. (eds.) *The Genetics of the Horse*. Wallingford, Oxfordshire: CABI Publishing. pp 392-393.

Utami, D.W. et al. (2017) Pencarian alel untuk identifikasi gen ketahanan penyakit hawar daun bakteri, Xa7 pada plasma nutfah padi lokal Indonesia. *Jurnal AgroBiogen*, 6(1):1-9. Doi:10.21082/jbio.v6n1.2010.p1-9.

Villa, T.C.C. et al. (2006) Defining and identifying crop landraces. *Plant Genetic Resources*. 3(3):373-384. Doi:10.1079/PGR200591.

Wood, D. & Lenné, J.M. (1997) The conservation of agrobiodiversity on-farm: questioning the emerging paradigm. *Biodiversity and Conservation*, 6(1):109-129. Doi:10.1023/A:1018331800939.

Worldatlas (2017) *The countries producing the most rice in the world*. Available from: https://www.worldatlas.com/articles/the-countries-producing-the-most-rice-in-the-world.html

DIGITAL OBJECT IDENTIFIERS (DOI) FOR LOCAL RICE GENETIC RESOURCES OF CENTRAL KALIMANTAN

Susilawati and Muhamad Sabran

INTRODUCTION

Central Kalimantan which consists of 13 regencies and one city has two main agro ecosystems, namely dry land (including rainfed) and swamp land (both tidal, swampy and peat land). There are many diverse local rice genetic resources found in this area. The characteristics of the rice plants are very high, high and medium. The grains are slim, small, round, long, hairy, grain color reddish, golden or blackish, etc. Responses to biotic and abiotic stresses are resistant to stem borer, blast disease, inundation, shade tolerant and tolerant to certain acidity levels. Local rice is generally 5-6 months old, and is widely planted by local people for consumption, seed production and traditional ceremonies (Susilawati et al. 2014). They are widely developed and utilized. Several rice accessions were also threatened with extinction and some are no longer be found. More than 170 local rice accessions from Central Kalimantan have been explored. Some have been characterized and collected at Genebank of Indonesian Center for Agricultural

Biotechnology and Genetic Resources (ICABIOGRAD) and Assessment Institute for Agricultural Technology (AIAT) of Central Kalimantan (Susilawati et al. 2017). Some of local rice accessions also have been registered by Center for Plant Variety Protection and Agricultural Licensing (PPVTPP).

In the management of genetic resources in AIAT of Central Kalimantan, the system of documentation and information on explored and collected local rice genetic resources are still limited to passport data (data relating to the origin of germplasm material) and characterization data (data relating to morphological and agronomy characters), as well as information on resilience and/or tolerance to biotic and abiotic stresses, which are compiled based on field information. At present the information system has been developed with a computer-based Decision Support System (DSS). In this system information on genetic resources that are explored and collected can be known for their distribution and existence, and equipped with cultivation information and recommendations for development and utilization and their level of suitability (Bhermana and Susilawati 2018). Local rice data information systems arranged in the form of a database system will be good and may ease the process in entering data into storage media, data validation during the entry stage, flexibility in the operation of data entry and access, data availability is always up to date and allows for rapid exchange of data (Kurniawan et al. 2004).

The development of food crop germplasm data base using the Digital Object Identifiers (DOI) system was introduced. The DOI has been adopted by the Scientific Advisory Committee (SAC) on the Global Information System (GLIS) and applied in the management and exchange of Plant Genetic Resources for Food and Agriculture (PGRFA). For the trials in Indonesia, local rice commodities from Central Kalimantan are also used as model plants. Implementation of the DOI system that is standard ISO 26324 would be application of all genetic resources of food crops (PGRFA) commodities. Indonesia, which is trusted as a coordinator for software development (GLIS-Toolkit) in order to support the implementation of the DOI system, has collaborated with the ITPGRFA-FAO Secretariat in Rome, Italy and the International Rice Research Institute (IRRI) in the Philippines to conduct dissemination and technical training to several ITPGRFA-FAO signatory countries.

This paper is a review of some of the results in the Central Kalimantan AIAT management of local rice genetic resources, which aims to provide information on the existence and availability of its local rice germplasm that can be exchanged and utilized in scientific development, through registration in the DOI system.

Exploration and Characterization of Local Rice Genetic Resource of Central Kalimantan

Exploration of food crops, especially local rice, has been carried out since 2013 and continues to be expanded. In 2013 exploration of genetic resources was carried out in the yard, then since 2014 exploration of food crops, especially local rice, has been carried out on various land agroecosystems, both in swamp land, such as tidal, swampy and peatland and dryland including rainfed land (Susilawati et al. 2015). Until April 2019 as many as 171 local rice accessions from Central Kalimantan have been explored, they consist 35 accessions (2013), 43 accessions (2014), 32 accessions (2015), 20 accessions (2016), 12 accessions (2017), 10 accessions (2018) and 19 accessions (2019). Based on its distribution, around 63% of accession is explored from dryland rice and 37% is from swamp land (Susilawati et al. 2018). Most of the exploration results have been collected at Genebank of ICABIOGRAD, and some are collected simply in Central Kalimantan AIAT.

Characterization performed on collected local rice is still limited to morphological or agronomic traits, which consists of leaf surface, flag leaf position, leaf neck color, leaf ear color, leaf book color, leaf strand color, leaf midrib color, leaf tongue color, leaf length, leaf width, stem segment color, stiffness of the stem, panicle type, pan panicle, panicle length, sterile lemma, lemma and palea color, sterile lemma color, color of grain tip, grain edge feather, color of the end of grain, color pistil head, class variety, flowering age, harvest age, number of tillers, plant height, habitus and loss (FAO 1996; IRRI 1996). Evaluation on amylose content, resistance to bacterial leaf blight, resistance to white-striped leaf pests, resistance to pests, resistance to aluminum poisoning and resistance to drought, were not carried out. This is due to the limitation of human resources to manage genetic resources and other facilities contained in Central Kalimantan AIAT.

Besides having specific characters of varieties, some local rice accessions that are explored and collected, have good adaptability to sub-optimal land. Rice accession explored from dryland agroecosystems, with the main constraints of low soil pH and high Al content (Kochian 1995), proved to be able to withstand these conditions. Even though it is known that high Al solubility can inhibit root growth and function, and therefore, the condition can reduce production between 25%-85% (Herrera and Estrella 2003). Two dryland local rice varieties known as Sentang variety from the Kapuas district and Sahuwi from the district Lamandau have a mechanism to defend themselves in these clenched conditions by extending their roots to soil pH 4.0 while the condition of the land is quite dry.

In the case of swamp agroecosystem, two varieties known as Siam Epang from Kotawaringin Timur district and Brenti from Pulang Pisau district may survive in the conditions of Fe poisoning by an excluder mechanism, which is a mechanism by which plants are able to accumulate excessive Fe^{2+} ions at the root, and inhibit the entry of excessive Fe^{2+} ions into the root zone (Marschner 1995). In addition, in the inundated conditions, all plants are also able to survive by forming new roots as their adaptability mechanism. This condition is in accordance with Yang et al. (2005), which states that the excluder mechanism contained in plants is related to ethylene production in roots. Ethylene in the form of small and gas-shaped molecules has a function to help the development of aerenchyma in the roots and also the formation of new roots as a way of adapting plants to stagnant conditions.

Management of Local Rice Genetic Resources Specific Location

The management system of genetic resources, especially local rice in Central Kalimantan AIAT has been carried out, among others, by compiling passport data (data relating to the origin of germplasm material), characterization data (data relating to morphological and agronomic characters) and information on resilience/tolerance to biotic and abiotic stresses, which are compiled based on field information. In information based on passport data, specifically for local rice, data is divided into local rice data specific to swamps and dryland, as shown in Table 1 and Table 2.

Table 1. Local rice genetic resources specific on swamp land of Central Kalimantan.

No.	Accesion number	Accesion name	Village	Sub Distric	District	Harvest time	Rice type
1	0041	Brenti	Tahai Jaya	Maliku	Pulang Pisau	130	White Rice
2	0042	Buntut kuda	Tahai Jaya	Maliku	Pulang Pisau	130	White Rice
3	0043	Kahayan	Tahai Jaya	Maliku	Pulang Pisau	130	White Rice
4	0044	Karang Dukuh	Palingkau	Dadahub	Kapuas	125	White Rice
5	0045	Kencana	Tahai Jaya	Maliku	Pulang Pisau	130	White Rice
6	0046	Lemo	Tahai Jaya	Maliku	Pulang Pisau	130	White Rice
7	0047	Pentet,	Tahai Jaya	Maliku	Pulang Pisau	130	White Rice
8	0048	Pikat	Tahai Jaya	Maliku	Pulang Pisau	130	White Rice
9	0049	Siam Banjar	Palingkau	Dadahub	Kapuas	125	White Rice
10	00410	Siam Epang	Palingkau	Dadahub	Kapuas	125	White Rice
11	00411	Siam Landak	Palingkau	Dadahub	Kapuas	125	White Rice
12	00412	Siam Lantik	Palingkau	Dadahub	Kapuas	125	White Rice
13	00413	Siam Mutiara	Dadahup	Dadahub	Kapuas	125	White Rice
14	00414	Siam Pedang	Palingkau	Dadahub	Kapuas	125	White Rice
15	00415	Siam Pudak	Palingkau	Dadahub	Kapuas	125	White Rice
16	00416	Siam Sekonyer	Kumai	Kumai	Arut Selatan	125	White Rice
17	00417	Siam Unus	Palingkau	Dadahub	Kapuas	125	White Rice
18	00418	Bajang	Palingkau	Dadahub	Kapuas	125	White Rice
19	00419	Balimau	Buntut Bali	Tewang SG	Katingan	125	White Rice
20	00420	Banyu Bilis	Buntut Bali	Tewang SG	Katingan	125	White Rice
21	00421	Barinsai	Palingkau	Dadahub	Kapuas	125	White Rice
22	00422	Behas bahandang	Jabiren	Jabiren Raya	Pulang Pisau	125	Red Rice
23	00423	Bulan	Palingkau	Dadahub	Kapuas	125	White Rice
24	00424	Buntut Raja	Palingkau	Dadahub	Kapuas	125	White Rice
25	00425	Cirendah	Buntut Bali	Tewang SG	Katingan	125	White Rice
26	00426	Gadabung	Timpah	Timpah	Kapuas	125	White Rice

At present the a computer-based information system has been developed by *Decision Support System* (DSS). In this system, information on explored and collected genetic resources is equipped with a map of whereabouts and their distribution. In addition, it is equipped with cultivation information and recommendations for its development and utilization (Bhermana and Susilawati 2018). PC-based Microsoft Access applications used in developing database systems with DSS are built as planning databases that are used for logical data structures in helping decision makers for planning and management (Hirouchi and Kosaka 1984). This application is also effective for database management systems (DBMS) (Oktarina and Kuswantoro 2011).

Table 2. Local rice genetic resources specific on upland of Central Kalimantan.

No.	No accesion	Name accesion	District	Status accesion	Old of Harvest	Rice type
1	0151	Garagai	Kapuas	Lokal	130	White Rice
2	0152	Humbang Inai	Katingan	Lokal	130	White Rice
3	0153	Kalanis	Katingan	Lokal	130	White Rice
4	0154	Kowong	Katingan	Lokal	130	White Rice
5	0155	Kumpang Emas	Lamandau	Lokal	130	White Rice
6	0156	lokal Gilai	Lamandau	Lokal	130	White Rice
7	0157	Lungkuh Tamuun	Barut	Lokal	130	White Rice
8	0158	Manyahi	Kapuas	Lokal	130	White Rice
9	0159	Mungkolandik	Gunung Mas	Lokal	130	White Rice
10	01510	Pahakung	Lamandau	Lokal	130	White Rice
11	01511	Siam Gunung	Barut	Lokal	130	White Rice
12	01512	Sibung rendah,	Katingan	Lokal	130	White Rice
13	01513	Talun Bajang	Barito Utara	Lokal	130	White Rice
14	01513	Talun Bitik	Barito Utara	Lokal	130	White Rice
15	01514	Tambangan	Barito Utara	Lokal	130	White Rice
16	01515	Tampui	Lamandau	Lokal	130	White Rice
17	01516	Umbang putih,	Katingan	Lokal	130	White Rice
18	01517	Buyung	Katingan	Lokal	130	White Rice
19	01518	Garagai	Kapuas	Lokal	130	White Rice
20	01519	Garu	Barito Utara	Lokal	130	White Rice
21	01520	Gilai	Lamandau	Lokal	130	White Rice
22	01521	Goyang pinggul	Barito Utara	Lokal	130	White Rice
23	01522	Humbang Inai	Lamandau	Lokal	130	White Rice
24	01523	Kalanis	Katingan	Lokal	130	White Rice
25	01524	Karang Dukuh	Kapuas	Lokal	130	White Rice
26	01525	Karundang	Kapuas	Lokal	130	White Rice

Regarding the availability of information maps, the application of geographic information systems (GIS) called Q-GIS 3.4 is also used to supplement databases in spatial format. In addition to facilitating, improving, storing, and displaying data, the application can also be used to update data and utilization in the future (Eswaran et al. 1992; Goodchildet al. 1992; Bhermana et al. 2002; Son and Rajendra 2008). As part of the GIS, a basic reconnaissance-scale land system map of land systems is also used to identify land suitability for each local rice plant genetic resources (PGR) in accordance with its habitat. This is beneficial for the development of in situ and *ex situ* according to the similarity of landforms, soil, vegetation, and relative climate uniform (FAO 1993; Suharta 2007).

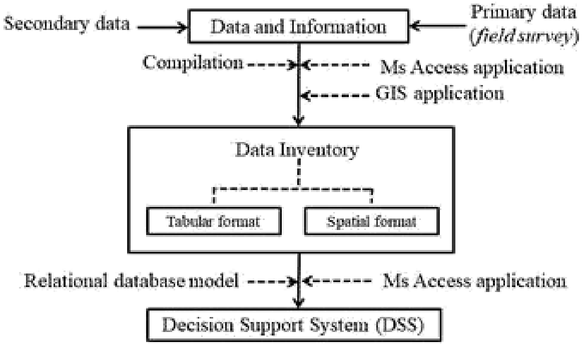


Figure 1. Diagram of DSS preparation procedure.

DOI Registration and Utilization of Central Kalimantan Local Plant Genetic Resources

DOI has been currently chosen for the mechanism of providing and identifying unique and permanent germplasm globally. In Indonesia the development of a food crop germplasm database with DOI system was introduced. This digital object identifier has also been adopted by the Scientific Advisory

Committee (SAC) on the Global Information System (GLIS) and applied in the management and exchange of Plant Genetic Resources for Food and Agriculture (PGRFA). Implementation of the DOI system that is standard ISO 26324 would be application to all PGRFA commodities. Indonesia, which is trusted as a coordinator for software development (GLIS-Toolkit) to support the implementation of the DOI system, has collaborated with the ITPGRFA-FAO Secretariat in Rome, Italy and the International Rice Research Institute (IRRI) in the Philippines to conduct dissemination and technical training to several ITPGRFA-FAO signatory countries. For the trials in Indonesia, local rice commodities from Central Kalimantan were used as model plants. This trial resulted in 96 local rice accessions from Central Kalimantan were registered with DOI (Table 3).

From a number of local rice accessions from Central Kalimantan, there is one local variety, namely Siam Mutiara from Kapuas district that has been used as an parental line in plant breeding The government has also released the R1 variety Dadahup,a progeny of Siam Mutiara. One other variety, Siam Epang from the East Kotawaringin district, has been bleached as a local superior variety. In addition, as many as five local varieties have been fingerprinted by ICABIOGRAD, and as many as 10 local rice varieties have been registered with the Center for Plant Variety Protection and Agricultural Licensing, with evidence in the form of a register of plant varieties (Susilawati et al. 2018). In 2019 there will be 10-15 registered local rice varieties.

Conclusion

1. The genetic resources of local rice in Central Kalimantan are spread in two main agroecosystems, namely dryland and swampland.

2. Management of local rice genetic resources that have been carried out is compiling passport data, characterization data, and information on resistance to biotic and abiotic stresses, as well as computer-based information systems by Decision Support System that is able to provide spatial information on the distribution of genetic resources, cultivation information and recommendations for compliance.
3. A total of 96 local rice accessions from Central Kalimantan have also been managed through the Digital Object Identifiers, so as to encourage the use of local rice PGR globally for the benefit of scientific knowledge.

Acknowledgments

Thank you to ICABIOGRAD and IRRI for facilitating the implementation of DOI registration, and financially supported for the holding of this conference. Thank you also to Dr. Andy Bhermana, SP., M.Sc. who has an active role in developing information systems for Central Kalimantan's agricultural resources based on spatial data.

REFERENCES

Bhermana, A., Hamdan J., Anyar A.R. & Peli M. (2002) Determination of Agricultural Land Regions Using Agroecological Zone (AEZ) Approach and Geographic Information System. A case study of Kotawaringin Barat Regency, Kalimantan, Indonesia. Proceeding of The Malaysian Society of Soil Science Conference 2002. *Application of Modern Tools in Agricultur*e, 36-39.

Bhermana, A. & Susilawati. (2018) Decision Support System Based On Database System of Genetic Resource For Central Kalimantan Local Crops to Develop *Ex Situ*s and In Situs. Naskah disampaikan pada The First International Conference on Genetic Resources and Biotechnology Bogor, August 20th – 21st 2018.

Eswaran H., Kimble J., Cook T. & Beinroth F. H. (1992) Soil diversity in the tropic: implications for agricultural development. *Myths and Science of Soils of the Tropics*, 29:1-16.

FAO (1993) *Guideline for Land-Use Planning*. FAO Development Series 1. Rome: Food and Agriculture Organization of The United Nation.

FAO (1996) *Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture*. Rome: Food and Agriculture Organization of The United Nation.

Kurniawan, H., Sutoro, Mamik, S., Tiur, S.S., Sri, G.B., Hadiatmi, Asadi, Nurwita, D., Sri, A., Rais, Ida, H., Somantri, Nani, Z., Minantyorini & Tintin, S. (2004) *Pengembangan Sistem Pangkalan Data (Database) Plasma Nutfah Tanaman Pangan*. Kumpulan Makalah Seminar 74 Hasil Penelitian BB-Biogen Tahun 2004. BB Biogen. Bogor.

Goodchild, M., Robert. H. & Stepen. W. (1992) Integrating GIS and spatial data analysis: problem and possibilities. *Int J Geographical Information Systems*, 6(5):407-423.

Herrera-Estrella, L. (2003) Use of Biotechnology od Increase Food Production on Acid Soid. Retrieved from http://europa.eu.int/ comm./reseach/conferences/2003/sad/proestrella-en.html.

Hirouchi, T & Kosaka, T. (1984) An effective database formation for decision support system. *Information Management*. 7(4):183-195.

IRRI (1996) *Standart Evaluation System for Rice.* 4th Edition. Philipina: INGER Genetic Resources Centre IRRI.

Kochian, L.V. (1995) Cellular mechanisms of aluminium toxicity and resistance in plant. *Annu Rev Plant Physiol Mol Biol,* 46:237-260.

Marschner, H. (1995*) Mineral Nutrition of Higher Plants. 2nd Ed.* Academic Press. New York: Harcourt Brace and Company.

Oktarina, N. & Kuswantoro, A. (2011) Pembelajaran berbasis IT aplikasi program MS access untuk meningkatkan pemahaman pada pokok bahasan inventarisasi. *Jurnal Penelitian Pendidikan,* 29(2):158-168.

Son, N.T. & Rajendra, P.S. (2008) GIS-assisted land evaluation for agricultural development in mekong delta, southern Vietnam. *Journal of Sustainable Development in Afrika*. 10(2):875-884.

Suharta, N. (2007) Sistem lahan barong tongkok di kalimantan: potensi, kendala, dan pengembangannya untuk pertanian lahan kering. *Jurnal Litbang Pertanian,* 26(1):1-8.

Susilawati, Saleh, M., Sintha, E., Sri, A. & Suparman. (2014) Inventarisasi sumberdaya genetik padi lokal spesifik Kalimantan Tengah. Prosiding Seminar Inovasi Teknologi Padi Mendukung Pertanian Bioindustri Tahun 2014. Balai Besar Penelitian Padi. Sukamandi.

Susilawati, Suparman, S., Agustini, Harmini, Sinta, E.P., Adrial & Massinai, R. (2015) Inventarisasi sumber daya genetik pertanian kalimantan tengah. *In*: Tri Puji P., Hakim K., Nurul H., Mia K. & Yadi S. (eds). Prosiding Seminar Nasional Sumber Daya Genetik Pertanian. p 326-338.

Susilawati, Rukayah, Sintha, E.P, Twenty, L., Munier, L.L., Bhermana, A. & Siahaan, M. (2017) Pengelolaan Sumber Daya Genetik Spesifik Kalimantan Tengah. Laporan Hasil Kegiatan Pengkajian Tahun 2017. BPTP Kalimantan Tengah. Palangka Raya.

Susilawati, Rukayah, Sintha, E.P, Twenty, L., Munier, L.L., Bhermana, A. & Siahaan, M. (2018) Pengelolaan Sumber Daya Genetik Spesifik Kalimantan Tengah. Laporan Hasil Kegiatan Pengkajian Tahun 2018. BPTP Kalimantan Tengah. Palangka Raya.

Yang, I.L., Zheng, S.J, Feng He. & Matsumoto H. (2005) Aluminium resistance requires resistance to acid stress: a case study with a spaniach that exudes oxalate rapidly when expose to Al stress. *Journal of Experimantal Botany,* 56(414):1197-1203.

THE FOUR YOGYAKARTA'S LOCAL RICE SUPERIORITY REGISTERED BY DIGITAL OBJECT IDENTIFIER

Setyorini Widyayanti, Kristamtini, and Nurul Hidayatun

INTRODUCTION

Indonesia is one of megadiversity country which has a diverse genetic resource: Rice is one of food crop genetic resources. Regional conditions (soil, rainfall type), Indonesia's topography and climate are very diverse, that contribute to rice accessions diversity which developed in each region of Indonesia. Some information states that rice is not an Indonesia native's plant. It is believed that rice have been first domesticated in Yang tse river valley, Hunan province, China in 5000-4000 BC and then referenced into Ganges river valley, which is around India and Sri Lanka, Korea, Japan and Southeast Asia (Indonesia, Malaysia, Philippines) (Vaughan et al. 2008; Gross and Zao 2014). In Indonesia, rice cultivation was spreading from west to eastern end sovereign part. It is not certainly known which one of Indonesia's territory that first made rice cultivation.

Rice does not consist only white rice, there is also rice accession which produces another colored rice, such as yellow, red, brown, purple and black rice (Sanghera et al. 2013). The color is due to anthocyanin pigment accumulated in the pericarp layer, seed shell or aleurone (Kristamtini 2014). Colored rice is a type of rice that has existed over long period. Colored rice is believed to have special features because it has a high nutrient content which can control various diseases and also it can extend longevity of life (Kushwaha 2016). Red rice has been consumed by the people in India, Sri Lanka and Bhutan on their special days and traditional ceremonies, whereas in China, black rice is only consumed by the royal family (king) hence black rice is also called forbidden rice, king's rice or prized rice (Sanghera et al. 2013; Oikawa et al. 2015; Kushawa 2016). In 2014, Yogyakarta Assessment Institute for Agricultural Technology (AIAT) has conducted an exploration, inventory and identified 76 Yogyakarta local rice accessions (Sudarmaji 2014; Kristamtini et al. 2015). Yogyakarta is one of the provinces in Indonesia.It is located in central Java, and cover an area of 3,185.80 km^{2}and It lies between 7°33'−8°12' South Latitude and 110°00'−110°50' East Longitude of Greenwichand stretches from 0-2,911 m above sea level (above sea level) (BPS 2014). Typology of Yogyakarta is quite diverse consisting of Mount Merapi area, lowland areas, coastal sandy land and mangrove area which contribute greatly to the existing of genetic resources diversity (Bappeda 2009). With the diversity of agro-ecosystem, it is possible that Yogyakarta have colored rice genetic resources.

The 76 local rice accessions consisted of white rice, red rice, black rice and sticky rice. Genetic resources management of AIAT consist of exploration, identification, characterization, evaluation and utilization so that it will not disappear or become extinct. Beside preserving and protecting genetic resources from extinction and biopiracy, the government has provided a legal law through a mechanism for registering both local varieties and breeding activity with the enactment of Law No. 29 of 2000 concerning Protection of Plant Varieties, government law No. 13

of 2004 concerning Registration of Plant Varieties and the Ministry of Agriculture Decree No. 67/2006 concerning the preservation and utilization of genetic resources (PPVTPP 2015).

An effort to improve conservation and sustainable use of food and agricultural genetic resources were also made by FAO (Food and Agriculture Organization) which is a United Nation organization. The exploration to identification activities of AIAT Yogyakarta were funded by FAO's Benefit Sharing Fund under the International Treaty of Plant Genetic Resources for Food and Agriculture (ITPGRFA), multicountry activities. This kind of platform is expected to reduce the information gap between gene bank curators, rice breeders and researchers who are involved in agricultural cultivation for supporting transfer of technology and information of genetic resources for developing a new variety. The platform also supports applied germplasm curation and forward-looking rice breeding programs and strategic rice research.

Sharing information regarding to plant genetic resources, especially annex-1 group of plants, has been carried out by countries who already signed the ITPGRFA treaty using a diverse data storage application system. To provide an easy genetic resources data information access between those countries, a digital system identification was needed. Multi-country construction of a test platform for development and allocation of unique identifiers for rice germplasm is a research activity coordinate by Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD), which make an effort to adopt and implement digital object identifier (DOI) as a global, unique, no-bias identities and develop a platform to establish automatized system to system connection to add value to the material being transferred within and from the Multilateral System (MLS), thus meeting both scientific needs and legal obligation of the Standard Material Transfer Agreement (SMTA).

The purpose of this paper is to provide information on the characteristics and superiority of four Yogyakarta colored rice cultivars which has been registered in a digital information system.

Digital Object Identifier (DOI) on Plant Genetic Resources

Indonesia is one of the countries which have ratified ITPGRFA in 2006. One consequence of ITPGRFA is that every country member agreed to establish a multilateral system of access and benefit sharing of plant genetic resource, which is efficient, effective, transparent, fair and in an equitable way. Instruments which can be used for information dissemination applications are Digital Object Identifier (DOI).

DOI is a permanent instrument that can be used to recognize or mark an electronic document. DOI is generally a series of unique and fixed codes (persistence) consisting of different numbers and letters on each object (entity). DOI can be used on all electronic documents, and currently it has been used to identify genetic resources accessions.

Rice is one of the commodities listed in Annex 1 of ITPGRFA and are listed as the main food commodities in many countries. However, rice genetic resources collection is not yet fully accessible in various treaty signatory countries. This is allegedly caused by the lack of information systems that support data access. Therefore, the dissemination of rice genetic resources diversity through DOI implementation is expected to be a part of an information system that supports, eases and facilitates data access.

AIAT Yogyakarta has successfully exploring, invented and carried out 76 Yogyakarta local rice. From identification and characterization activity, until the end of 2014, there were 55 accession of Yogyakarta local rice that was ready to be registered in the DOI system. The results have been registered into ICABIOGRAD registration system.

The 55 accession of Yogyakarta local rice which uploaded in ICABIOGRAD registration system have a variety of features. A few of them are now rare or only be found in a specific location in Yogyakarta. Another specific feature is a distinguished pericarps color (white, red or black pericarp color). Other accessions provide a specific fragrant aroma. Some of it are local rice accessions which have a potential value such as high production. There were also accessions which are specific on irrigated rice, upland rice also sticky rice. Information related to Yogyakarta local rice digital numbering can be accessed through https://ssl.fao.org/glis/. Yogyakarta local rice which has been registered through the ICABIOGRAD digital numbering system is listed in local ID numbers 05020-30654 up to numbers 05020-30708 (Widyayanti et al. 2018).

Morphological Characters of Yogyakarta Local Rice Registered by DOI

AIAT Yogyakarta has already explored and carried out 76 Yogyakarta local rice accessions. 55 local rice accessions have been fully identified and registered in DOI through ICABIOGRAD registration system. Registered accessions consisted of 5 accessions of red rice, 5 accessions of black rice, 30 accessions of white rice, and 15 accessions of sticky rice.

According to Indonesian Ministry of Agriculture Decree, 4 Yogyakarta local rice accessions which have been assigned with

DOI have been released as a local superior variety. The four rice varieties are 1) Mandel Handayani which has been released into local superior variety through Indonesian Ministry of Agriculture Decree number 2227/Kpts/SR.120/5/2009, May 19, 2009; 2) Segreng Handayani which has been released into local superior variety through Indonesian Ministry of Agriculture Decree, number 2226/Kpts/SR.120/5/2009, May 19, 2009; 3) Cempo Merah Sleman through Indonesian Ministry of Agriculture Decree number 126/HK.540/C/03/2019, March 6, 2019 which was released as a local superior variety under the name of Sembada Merah and 4) Pari Ireng through Indonesian Ministry of Agriculture Decree number 125/HK.540/C/03/2019, dated March 6, 2019 which was released as a local superior variety under the name of Sembada Hitam.

Mandel Handayani and Segreng Handayani are specific local red rice variety from Gunungkidul. Gunungkidul is one of Yogyakarta's district area which mainly consists of dry land, and thus both rice varieties are commonly cultivated as an upland rice. Sembada Merah is red rice and Sembada Hitam is black rice, both of them are from Sleman, one of Yogyakarta's district area. Sembada Merah and Sembada Hitam are commonly cultivated as irrigated rice. Four of Yogyakarta colored rice are registered in ICABIOGRAD local ID number05020-30660; 05020-30666; 05020-30707 and 05020-30707 (Table 1).

Table 1. Identification of DOI, local ID, type and origin of 4 Yogyakarta colored rice varieties.

Accessions	DOI	Local ID	Type of rice	Area origin
Mandel Handayani	10.18730/SF1F	05020-30660	Upland rice	Gunungkidul, Yogyakarta
Segreng Handayani	10.18730/SF1N4	05020-30666	Upland rice	Gunungkidul, Yogyakarta
Sembada Merah	10.18730/SF2Y8	05020-30707	Irrigated rice	Sleman, Yogyakarta
Sembada Hitam	10.18730/SF2Z9	05020-30708	Irrigated rice	Sleman, Yogyakarta

Sources: Kristamtini et al. (2015); Widyayanti et al. (2018).

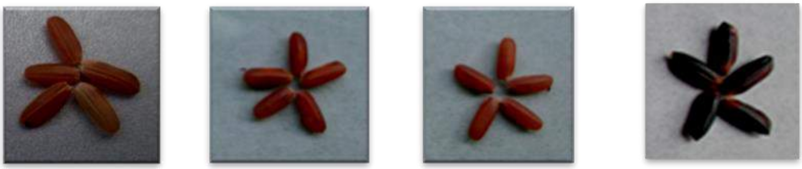
Descriptions of morphological characters of the 4 Yogyakarta specific colored rice varieties are presented in Table 2 and Figure 1. Although morphological characters observations are strongly influenced by the environmental conditions in which they growth, these observations are considered the most accurate for identifying the phenotypic performance of individuals or groups of plants (Kristamtini et al. 2015). According to morphological character, plant height of red rice from both Gunungkidul (Mandel Handayani and Segreng Handayani) and Sleman (Sembada Merah) has a medium-sized plant (110-130 cm for irrigated rice and 90-125 cm for upland rice) (Komnas Plasma Nutfah 2003). Similarly, Sembada Hitam has a medium plant height (125 cm). Medium-sized plant height is preferred by farmers, because of easier harvesting process.

Mandel Handayani grain seeds

Segreng Handayani grain seeds

Sembada Merah grain seeds

Sembada Hitam grain seeds



Mandel Handayani rice

Segreng Handayani rice

Sembada Merah rice

Sembada Hitam rice

Figure 1. Morphology performance of grain and rice seeds of 4 Yogyakarta colored rice.

Table 2. Description of the morphological character of 4 Yogyakarta specific colored rice varieties.

Characters	Accessions			
	Mandel Handayani	Segreng Handayani	Sembada Merah	Sembada Hitam
Days of maturity (day)	115	112	111	140
Plant height (cm)	141.00	93.50	91.00	125.00
Number of productive tillers	2 – 5	8	10 – 15	10 – 15
Number of full grainsper panicle	193	143	112	141
Yield potential (t/ha)	4.50	4.00	6.00	6.00
Weight of 1000 seeds (g)	27.00	24.33	26.67	24.00
Plant shape	slightly curved erect	erect	erect	erect
Feet color	green	purple	purple	green
Culm color	green	purple	green	green
Leaf blade color	colorless	colorless	purple lines	purple lines
Ligule color	colorless	colorless	puple lines	white
Flag leaf	semi-erect	erect	erect	erect
Grain shape	round	slender	slender	slender
Grain loss	easiest	easiest	Easy	moderate
Culm strenght	very weak	weak	strong	strong
Grain color	yellow	Reddish yellow	Straw yellow	Black with yellow lines
Pericarp color	Red to endosperm	Red	Red	Black

Sources: Kementerian Pertanian (2009; 2009a; 2019; 2019a).

Days to maturity is relatively similar between the red rice accessions (Mandel Handayani, Segreng Handayani and Sembada Merah) which is around average of 112 days so that it is classified as early maturity rice (105-124 days). Sembada Hitam is classified as medium-days to maturity rice (125-150 days) (BB Padi 2009). The number of productive tillers of Mandel

Handayani and Segreng Handayani is relatively small (<5-9 tillers). However, Mandel Handayani has greater number of productive tiller than Segreng Handayani. This is directly proportional that Mandel Handayani has higher yield potential than Segreng Handayani.

Pests and Diseases Resistance of Yogyakarta Local Superior Rice Varieties Registered by DOI

The main pests and diseases that usually attack rice cultivation in Indonesia are rice stem borer, brown planthopper, blast and bacterial leaf blight. Pests and diseases attack in all stages of growth in rice cultivation. If these pests and diseases were attacking in generative phase it could decline rice production.

Several efforts that could be considered effective for controlling pest and disease can be done through improved cultivation techniques, such as using healthy seeds, plant spacing, proper fertilization, environmental sanitation and prevention activity (Sudir et al. 2012). Other effort to prevent pests and diseases can be done by cultivating resistant varieties. Most of resistant varieties can be found in local varieties. Local varieties are believed to have resistance genes that can be used in breeding program.

The scoring observation showed that Mandel Handayani and Segreng Handayani were somewhat resistant of rice stem borer, sheath blight and bacterial leaf blight. The analysis using insect feeding behavior instrument namely electrical penetration graph (EPG) showed that Sembada Hitam followed by Sembada Merah have resistance to brown planthopper (Pemerintah Kabupaten Sleman 2018).

The resistance results analysis to bacterial leaf blight both through scoring intensity of disease progression and by using quantitative assessment indicators; namely AUDPC (Area Under Disease Progress Curve) with 2 weeks observation showed that Sembada Merah and Sembada Hitam AUDPC’s value are lower than Ciherang's national variety. This indicated that Sembada Merah and Sembada Hitam have resistance to bacterial leaf blight (Pemerintah Kabupaten Sleman 2018).

These results are in line with Sutoro et al. (2018), which state that Yogyakarta specific colored rice varieties contain resistance gene to blast and bacterial leaf blight. Sembada Hitam has a *Pii* resistance gene against blast and Xa7-LD40 gene against bacterial leaf blight. Sembada Merah has resistance gene (Xa4-LD6) against bacterial leaf blight.

Physical and Chemical Properties of Yogyakarta Specific Colored Rice Varieties Registered by DOI

Despite of having specific morphological characteristics, the four Yogyakarta specific colored rice varieties also have specific physical and chemical quality characters. The character of physical quality can be assessed through grain and milled rice.

The physical quality character of milled rice can be determined through the percentage of head rice, percentage of broken rice, brewers, chalkiness and biological damaged kernel (Indrasari et al. 2016), which in Indonesia must standardize as SNI No. 01-6128-2008 in 2008. According to this standard, head rice is a rice grain with a greater size than or equal to 0.75 parts of whole rice grains. Head rice is a component of physical quality of rice which directly affects the level of consumer acceptance. The results of laboratory analysis showed the percentage of head rice: Segreng Handayani (86.95%), Mandel Handayani (82.63%) and Sembada Merah (80.63%) and smaller Sembada Hitam (68.85%).

Table 3. The physical quality character of milled rice of 4 Yogyakarta specific colored rice varieties.

Physical quality of milled rice	Accessions			
	Mandel Handayani	Segreng Handayani	Sembada Merah	Sembada Hitam
Head rice (%)	82.63	86.95	80.63	68.85
Broken rice (%)	10.59	12.94	19.27	30.82
Brewers (%)	2.14	0.11	0.11	0.33
Chalkiness (%)	4.64	0.08	2.58	2.81
Damaged kernel (%)	0	0	0.47	0

Sources: Purwaningsih et al. 2008; BPTP Yogyakarta 2018; Pemkab Sleman 2018.

High percentage of broken rice, which occurs in mostly in Sembada Hitam, causes declining level of consumer acceptance, because consumers are less interested in the physical condition of rice that is not intact. The higher percentage of broken rice and brewers will decrease its economic value. Broken rice is a rice grain with has larger size than 0.2 to smaller than 0.8 parts of whole rice grains, and brewers is a rice grain with smaller size less than 0.2 parts of whole grains of rice. Mandel Handayani, Segreng Handayani, Sembada Merah and Sembada Hitam rice percentages still fulfill quality standards at all quality levels (SNI 2008).

Chalkiness is a brown rice which has white chalky and soft textured due to physiological factors, including greenish kernel (immature grains) which is harvested when the rice is not yet optimum. The four of Yogyakarta local superior rice varieties has chalkiness rice: Mandel Handayani (4.64%), Segreng Handayani (0.08%), Sembada Merah (2.58%) and Sembada Hitam (2.81%). Based on SNI 2008, chalkiness rice that meets the standards quality is Segreng Handayani, Sembada Merah and Sembada Hitam which have a maximum content of chalkiness of 3% of sample weight.

Biological damaged kernel is a damaged kernel caused by mechanical, physiological and pathological factors (SNI 2008; Febriandi 2017). The four of Yogyakarta specific colored rice varieties still meet the SNI 2008 quality standard of rice, which has the maximum standard grain damage of 3% of sample weight.

The chemical quality character of the four varieties is presented in table 4. Chemical quality characteristics are needed to determine the nutrient content that shows the quality of rice that can fulfill nutritional needs. The nutritional of rice which can indicate the level of glutinous rice is amylose content. The higher the amylose content, the lower the glutinous rice level. According to Damardjati (1988), the classification of amylose content is divided into waxy (1-2%), very low (2-9%), low (10-20%), intermediate (20-25%) and high (25-30%).

Sembada Hitam has the lowest amylose content (5.89%) so the rice texture is very glutinous but not waxy like sticky rice. Sembada Merah has 21.42% of amylose content so that the texture of rice is relatively glutinous. Mandel Handayani and Segreng Handayani are upland rice, with 28.77% and 29.55% amylose content, respectively, and are still classified as glutinous textured rice which tends to be rather hard.

Nutrient content that is a feature characteristic of colored rice is anthocyanin content. Anthocyanin is a phenolic compound that belongs to flavonoid group. The color difference formed in rice pericarp is influenced by the composition of the anthocyanin compounds. Anthocyanin functions as an antioxidant that has healthy benefit, such as preventing premature aging, protecting the stomach from damage, inhibiting tumor cell development, anti-inflammatory and anticancer compounds, preventing obesity and diabetes, improving brain memory, preventing neurological diseases and warding off free radicals (Kushwaha 2016).

Sembada Hitam's anthocyanin content showed the highest value (369.5µ/100g) compared to the red rice varieties. This is in line with the opinion of Sutharut and Sudarat (2012), which states that the highest anthocyanin content is found in black rice when compared to other colored anthocyanin rice or white rice.

Table 4. The chemical quality characteristics of milled rice of 4 Yogyakarta specific colored rice varieties.

Chemical quality parameters	Accessions			
	Mandel Handayani	Segreng Handayani	Sembada Merah	Sembada Hitam
Water content (%)	12.64	14.38	11.83	13.60
Ash content (%)	1.24	1.18	0.88	0.43
Protein (%)	10.2	9.25	9.04	5.51
Fat (%)	2.19	2.50	1.59	1.85
Crude fiber (%)	3.46	3.97	0.45	5.77
Amylose (%)	28.77	29.55	21.42	5.89
Amilopektin (%)	40.59	40.58	45.65	-
Starch (%)	69.36	70.03	67.07	89.05
Anthocyanin (cyanidin-3-glucosida µ/100g	nd	nd	1.2	369.5

Sources: Purwaningsih et al. 2008; Kristamtini et al. 2018; Pemerintah Kabupaten Sleman 2018; nd = not identified.

Concluding Remarks and Future Perspective

Four Yogyakarta specific colored rice varieties with its superiority in morphological characters, resistance to certain pests and diseases, physical and chemical quality characters have been assigned with the DOI system and thus, they can be included into Multi-Lateral System for ABS. Complete information is expected to be related to the diversity of genetic resources of rice, especially colored rice. Information related to Yogyakarta genetic resources diversity can be used as an ingredient of crossbreeding inmodern breeding programs.

REFERENCES

Bappeda D.I.Y. (2009) Penyusunan rencana pengelolaan plasma nutfah/sumber daya genetik tanaman di Daerah Istimewa Yogyakarta. Final Report. Kegiatan kerjasama PT. Gama Multi Usaha Madiri Yogyakarta dengan Badan Perencanaan Pembangunan Daerah Pemerintah Provinsi Daerah Istimewa.

BB Padi (2009) Pedum IP Padi 400; *Peningkatan Produksi Padi Melalui Pelaksanaan IP Padi 400*. Ministry of agriculture republic of Indonesia.

BPTP Yogyakarta (2018) Data hasil pengujian laboratorium analisis mutu giling beras.

BPS (2014) *Daerah Istimewa Yogyakarta dalam Angka*. Yogyakarta: Badan Pusat Statistik Daerah Istimewa Yogyakarta-Bappeda Daerah Istimewa Yogyakarta.

Damardjati, D.S. (1988) *Struktur Kandungan Gizi Beras, Padi, Buku 1*. Indonesian Agency for Agricultural Research and Development. Ministry of agriculture republic of Indonesia.

Febriandi, E. (2017) Karakterisasi sifat fisikokimia dan fungsional padi lokal Mayang Pandan dari Bangka Belitung. IPB University.

Gross, B.L & Zhao, Z. (2014) Archaeological and genetic insight into the origins of domesticated rice. PNAS, 111(17):6190-9197.

Indrasari, S.D., Rakhmi, A.T., Subekti, A. & Kristamtini (2016) Mutu fisik, mutu giling dan mutu fungsional beras varietas lokal Kalimantan Barat. *Jurnal Penelitian Pertanian Tanaman Pangan*, 35(1):19-28.

Kementerian Pertanian (2009) Pelepasan padi gogo beras merah lokal mandel sebagai varietas unggul dengan nama mandel handayani. SK Menteri Pertanian nomor 2227/Kpts/SR.120/5/ 2009 tanggal 19 Mei 2009. Jakarta: Ministry of Agriculture Republic of Indonesia.

Kementerian Pertanian (2009a) Pelepasan padi gogo beras merah lokal Segreng sebagai varietas unggul dengan nama Segreng Handayani. SK Menteri Pertanian nomor 2226/Kpts/SR.120/5/ 2009 tanggal 19 Mei 2009. Jakarta: Ministry of Agriculture Republic of Indonesia.

Kementerian Pertanian (2019) Pelepasan galur padi sawah lokal Sembada Hitam sebagai varietas unggul dengan nama Sembada Hitam. SK Menteri Pertanian nomor 125/HK.540/C/ 03/2019 tanggal 6 Maret 2019. Jakarta: Ministry of Agriculture Republic of Indonesia.

Kementerian Pertanian (2019a) Pelepasan galur padi sawah lokal Sembada Merah sebagai varietas unggul dengan nama Sembada Merah. SK Menteri Pertanian nomor 126/HK.540/C/ 03/2019 tanggal 6 Maret 2019. Jakarta: Ministry of Agriculture Republic of Indonesia.

Komnas Plasma Nutfah (2003) Panduan sistem karakterisasi dan evaluasi tanaman padi. Translated by Silitonga, T.S., Somantri, I.H., Daradjat, A.A. & Kurniawan, H. Indonesian Agency for Agricultural Research and Development.

Kristamtini, Wirasti., C.A., Widyayanti, S., Indrasari, S.D., Purwaningsih, H., Wiranti, E.W. & Yolanda, K. (2018) Laporan Akhir Tahun 2018. Kegiatan Pengelolaan Sumber Daya Genetik Spesifik D.I. Yogyakarta. Ministry of Agriculture Republic of Indonesia.

Kristamtini, Sutaryo, B., Wiranti, E.W., Pustika, A.B., Indrasari, S.D., Wirasti, C.A. & Hatmi, R.U. (2015) Laporan Akhir Tahun

2015. Kegiatan Pengelolaan Sumber Daya Genetik Spesifik D.I. Yogyakarta. Indonesian Agency for Agricultural Research and Development Ministry of Agriculture Republic of Indonesia.

Kristamtini (2014) Kajian Genetik Warna Beras. Disertasi. Program Pascasarjana Fakultas Pertanian Universitas Gadjah Mada. (Unpublished).

Kushwaha, U.K.S. (2016) *Black Rice*. Switzerland: Springer International Publishing.

Oikawa, T., Maeda, H., Oguchi, T., Yamaguchi, T., Tanabe, N., Ebana, K., Yano, M., Ebinita, T. & Izawa, T. (2015) The birth of a black rice gene and its local spread bay introgression. *The Plant Cell Online Journal*, ORCID ID: 0000-0002-5276-1610 (T.E).

Pemerintah Kabupaten Sleman (2018) Proposal usulan pelepasan padi beras merah Sembada Merah dan padi beras hitam Sembada Hitam. Presented on sidang pelepasan varietas, Pusat Penelitian dan Pengembangan Tanaman Pangan, Bogor.

PPVTPP (2015) Pusat PVTPP Giat Sosialisasikan Pentingnya Peran Perlindungan dan Pendaftaran Varietas Tanaman dalam Pemuliaan Tanaman. Retrieved from http://pvtpp.setjen.pertanian.go.id/.

Purwaningsih, H., Kristamtini, & Widyayanti, S. (2008) Mutu fisik dan organoleptik padi merah varietas lokal (Cempo Merah, Mandel dan Segreng) sebagai plasma nutfah padi propinsi D.I. Yogyakarta. Prosiding Seminar Nasional Inovasi Teknologi dan Mengantisipasi Perubahan Iklim Global Mendukung Ketahanan Pangan. Buku 4. Balai Besar Penelitian Tanaman Padi. Indonesian Agency for Agricultural Research and Development. p. 1481-1488.

Sanghera, G.S., Kashyap, S.C. & Parray, G.A. (2013) Genetic variation for grain yield and related traits in temperate red rice

(Oryza sativa L.) Ecotypes. *Notulae Scientia Biologicae*, 5(3):400-406.

SNI (2008) *Standar Nasional Indonesia: Beras*. SNI 6128:2008. Badan Standardisasi Nasional.

Sudarmaji (2014) Pengelolaan sumberdaya genetic spesifik lokasi Daerah Istimewa Yogyakarta. Annual Report. Yogyakarta: Balai Pengkajian Teknologi Pertanian Yogyakarta.

Sudir, B. Nuryanto, T.S. Kadir. 2012. Epidemiologi, patotipe dan strategi pengendalian penyakit hawar daun bakteri pada tanaman padi. *Iptek Tanaman Pangan*, 7(2):79-87.

Sutharut, J. & Sudarat, J. (2012) Total anthocyanin content and antioxidant activity of germinated colored rice. *International Food Research Journal*, 10(1):215-221.

Sutoro, Utami, D.W., Hidayatun, N., Widowati, L.R., Sabran, M., Priyatno, T., Koswanudin, D., Risliawati, A., Azizah, N. & Kurniawan, H. (2018) Sidik Jari DNA dan Allele Mining untuk Perlindungan dan Pemanfaatan Sumber Daya Genetik Padi Lokal. Laporan Kegiatan Akhir KP4S.BB-BIOGEN. p. 28 (Unpublished).

Vaughan, D.A., Lu, B.R. & Tomooka, N. (2008) Was Asian rice (*Oryza sativa*) domesticated more than once? *Rice,* 1:16-24. Doi:10.1007/s12284-008-9000-0.

Widyayanti, S., Kristamtini & Sudarmaji (2018) Digital Object Identifier (DOI) for rice germplasm collection at AIAT Yogyakarta. In International Conference on Genetic Resources for Effective Crop Improvement. Ministry of Agriculture. Indonesian Agency for Agricultural Research and Development. Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development. Bogor, Indonesia 20 – 21 Agustus 2019.

DOI IMPLEMENTATION ON RICE GERMPLASMS OF BRRI GENEBANK: PRESENT AND FUTURE ASPECTS

Armin Bhuiya, Md. Shahjahan Kabir, and Md. Khalequzzaman

INTRODUCTION

Global warming and climate change is harshly affecting biodiversity around the world. Bangladesh is blessed with diverse economically important plant genetic resources as it belongs within the geographical domain of Indian mega center of origin and also located near the Chinese mega center of origin (Hossain et al. 2014). These resources have high potential genetic variations, which can provide tremendous contribution in plant breeding to overcome various challenges by developing new varieties with high yield, biotic and abiotic stress tolerance etc.

Sharing or exchange of plant genetic resources (PGRs) has gained high attention worldwide to address certain environmental challenges, mitigate global hunger and achieve food security. Different organizations have different locally assigned identifiers for their PGRs, which can create confusions when these PGRs are transferred across organizations. Such problem can be minimized by a globally unique, persistent identifier such as Digital Object Identifier (DOI), which should maintain consistency over time, and provide proper recognition of rights and obligations as well as facilitating access to research outcomes produced by subsequent recipients of the materials (Alercia et al. 2018). Though DOI is only very recently introduced in Bangladesh, BRRI Genebank has pioneered in the registration of germplasms to obtain DOI in GLIS.

Background

International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) of the United Nations (FAO) is facilitating research, conservation and exchange of germplasms worldwide through the development and promotion of the use of Digital Object Identifiers (DOIs), which is an international standard to identify plant germplasm worldwide. The Treaty defines PGRFA as any genetic material of plant origin of actual or potential value for food and agriculture. This broad definition encompasses not only accessions conserved in genebanks and PGRFA conserved *in situ*, but also breeding lines, research materials, and protected modern varieties (Alercia et al. 2018).

Bangladesh has several National Biodiversity strategies and action plans for the conservation and sustainable use of plant genetic resources for food security, and modern agricultural practice like National Agriculture Policy-2018, National Biodiversity Strategy and Action Plan of Bangladesh 2016 to 2021, Access and beneficiary policy, Biosafety Guidelines of Bangladesh-2007, National Biosafety Framework-2007 and 7th Five Year Plan. It already has collaboration with FAO, IRRI, CIMMYT, ICRISAT, SAARC, AFACI, CIP, ICARDA and AVRDC for germplasm exchange and plant improvement for climate resilience. Bangladesh also has SMTA with AFACI and CGIAR institutes (Salam 2018).

Genebanks worldwide are conserving millions of accessions, some of which are duplicates but most carry valuable information associated with them. Sometimes this information is lost during the transfer of the material from one holder to another. Different user communities such as plant breeders, data curators, researchers and extension officers follow different methods to assign identifiers. This lack of standardization complicates the exchange of PGRFA data worldwide and creates challenges for effective conservation and sustainable use of PGRFA. Several communities (genebank community, genomics community, plant breeders, journal editors etc.) have emphasized the importance of creating and adopting Permanent Unique Identifiers for improved identification of PGRFA for better collaboration for conservation, research and breeding. By using the DOI standards adopted by the Governing body of the International Treaty, users will be able to identify and document their plant material uniquely and permanently and this will facilitate data interoperability among different systems. DOIs can be used to identify PGRFA held by any individual or organizations including genebanks, plant breeders, geneticists, other plant scientists, extension officers, seed companies, plant variety protection offices, gardeners, farmers, landowners and land managers (Alercia et al. 2018).

According to Alercia et al. 2018, DOI system brings the following new opportunities for users:

1. It exposes the material to the public or collaborators in a format that can be resolved by humans as well as computers
2. It enables information on the materials to be harvested by robots searching publications and online databases that refer to the PGRFA by its DOI and thus making them more readily available

3. It facilitates access to the information about the PGRFA and related PGRFA by pointing to websites and systems where detailed information is created, maintained and made available to the public
4. It provides a simple way of for recipients of material under an SMTA to comply with their obligations under the provisions of Article 6.9[3]of the SMTA, simply by using the DOI in their publications and online datasets to refer to the material received
5. It helps developers of PGRFA to maintain their records and comply with their obligations under SMTA article 6.5[b]
6. It enables families of PGRFA to be identified and thus jointly searched. For example, it enables a genebank manager to easily find all publications and online datasets created by recipients of accessions from the genebank
7. It provides a simple reliable mechanism to identify accessions that are duplicated across genebank
8. It facilitates interoperability between databases, by providing a single common standard for sample identification used by all communities
9. It enables collaborating laboratories, should they wish to track samples between them with any appropriate degree of precision while each laboratory continues to use its own in-house sample tracking system, providing assurance that they are working on the same material.

DOI Application in Bangladesh

Bangladesh Rice Research Institute (BRRI) has pioneered the registration of DOI for germplasms in Bangladesh. BRRI has 8578 germplasms registered as accession conserved in genebank (Table 1). A total of seventy germplasms were registered during the training program of "The 2nd International Workshop and Training on Software Testing to Support DOI Implementation" held at ICABIOGRAD, Indonesia on 26/04/19 to 2/05/2019 (Table 2). These are BRRI developed HYVs with special features. Most of these varieties are not only in high demand at farmers' field, but also have frequent utilizations and incorporation in new breeding programs (Table 3).

Table 1. Present status of BRRI genebank collection.

Variety/Line	Accession (no.)
Indigenous *indica* (Local)	5373
Indigenous *indica* (Pure Line)	857
Local *indica* varieties and breeding lines	565
Exotic *indica* (IRRI, China, USA, Turkey etc.)	1610
Japonica (Korea, Japan, Taiwan, N. China)	103
Wild rice (4 species)	46
Unknown	24
Total	8578

(Source: GRS Division, BRRI, 2019).

Table 2. List of BRRI germplasms registered in the GLIS.

Sl.No.	WIEWS	PID	Genus	NAME	Acc. No.	DOI
1	BGD002	00BA47	Oryza	BR-1	6865	10.18730/SBP66
2	BGD002	00BA47	Oryza	BR-3	4759	10.18730/SBP77
3	BGD002	00BA47	Oryza	BR-10	7089	10.18730/SBP88
4	BGD002	00BA47	Oryza	BR-11	6707	10.18730/SBP99
5	BGD002	00BA47	Oryza	BR-14	6872	10.18730/SBPAA
6	BGD002	00BA47	Oryza	BR-16	6874	10.18730/SBPBB
7	BGD002	00BA47	Oryza	BR-21	6199	10.18730/SBPCC
8	BGD002	00BA47	Oryza	BR-22	7090	10.18730/SBPDD
9	BGD002	00BA47	Oryza	BR-23	7091	10.18730/SBPEE
10	BGD002	00BA47	Oryza	BR-24	4276	10.18730/SBPFF

Table 2. Continue.

Sl.No.	WIEWS	PID	Genus	NAME	Acc. No.	DOI
11	BGD002	00BA47	Oryza	BR-25	4277	10.18730/SBPGG
12	BGD002	00BA47	Oryza	BR-26	4278	10.18730/SBPHH
13	BGD002	00BA47	Oryza	BRRI dhan27	4408	10.18730/SBPJJ
14	BGD002	00BA47	Oryza	BRRI dhan28	4409	10.18730/SBPKK
15	BGD002	00BA47	Oryza	BRRI dhan29	4410	10.18730/SBPMM
16	BGD002	00BA47	Oryza	BRRI dhan30	4411	10.18730/SBPNN
17	BGD002	00BA47	Oryza	BRRI dhan31	4412	10.18730/SBPPP
18	BGD002	00BA47	Oryza	BRRI dhan32	4413	10.18730/SBPQQ
19	BGD002	00BA47	Oryza	BRRI dhan33	7092	10.18730/SBPRR
20	BGD002	00BA47	Oryza	BRRI dhan34	7093	10.18730/SBPSS
21	BGD002	00BA47	Oryza	BRRI dhan36	6879	10.18730/SBPTT
22	BGD002	00BA47	Oryza	BRRI dhan37	7094	10.18730/SBPVV
23	BGD002	00BA47	Oryza	BRRI dhan38	7095	10.18730/SBPWW
24	BGD002	00BA47	Oryza	BRRI dhan39	7096	10.18730/SBPXX
25	BGD002	00BA47	Oryza	BRRI dhan40	7097	10.18730/SBPYY
26	BGD002	00BA47	Oryza	BRRI dhan41	7098	10.18730/SBPZZ
27	BGD002	00BA47	Oryza	BRRI dhan42	6214	10.18730/SBQ0*
28	BGD002	00BA47	Oryza	BRRI dhan43	6215	10.18730/SBQ1~
29	BGD002	00BA47	Oryza	BRRI dhan44	7099	10.18730/SBQ2$
30	BGD002	00BA47	Oryza	BRRI dhan45	6880	10.18730/SBQ3=
31	BGD002	00BA47	Oryza	BRRI dhan46	7100	10.18730/SBQ4U
32	BGD002	00BA47	Oryza	BRRI dhan47	6881	10.18730/SBQ50
33	BGD002	00BA47	Oryza	BRRI dhan48	7980	10.18730/SBQ61
34	BGD002	00BA47	Oryza	BRRI dhan49	7101	10.18730/SBQ72
35	BGD002	00BA47	Oryza	BRRI dhan50	6882	10.18730/SBQ83
36	BGD002	00BA47	Oryza	BRRI dhan51	7319	10.18730/SBQ94
37	BGD002	00BA47	Oryza	BRRI dhan52	7320	10.18730/SBQA5
38	BGD002	00BA47	Oryza	BRRI dhan53	7321	10.18730/SBQB6
39	BGD002	00BA47	Oryza	BRRI dhan54	7322	10.18730/SBQC7
40	BGD002	00BA47	Oryza	BRRI dhan55	7323	10.18730/SBQD8
41	BGD002	00BA47	Oryza	BRRI dhan56	7324	10.18730/SBQE9
42	BGD002	00BA47	Oryza	BRRI dhan57	7325	10.18730/SBQFA
43	BGD002	00BA47	Oryza	BRRI dhan58	8001	10.18730/SBQGB
44	BGD002	00BA47	Oryza	BRRI dhan59	8002	10.18730/SBQHC
45	BGD002	00BA47	Oryza	BRRI dhan61	8004	10.18730/SBQJD
46	BGD002	00BA47	Oryza	BRRI dhan62	8005	10.18730/SBQKE
47	BGD002	00BA47	Oryza	BRRI dhan63	8006	10.18730/SBQMF
48	BGD002	00BA47	Oryza	BRRI dhan64	8007	10.18730/SBQNG