



The contribution of the International Rice Genebank to varietal improvement and crop productivity in Eastern India

Donald Villanueva¹ · Melinda Smale² · Nelissa Jamora³ · Grace Lee Capilit¹ · Ruairaidh Sackville Hamilton¹

Received: 8 February 2020 / Accepted: 22 April 2020 / Published online: 2 June 2020
© The Author(s) 2020, corrected publication 2021

Abstract

Using survey dataset collected from nearly 9000 farmers along with pedigree and evaluation data, this study measures the contribution of the International Rice Genebank (IRG) to varietal improvement and rice productivity of farmers in Eastern India. We empirically test the relationship of ancestry to productivity changes while controlling for the effects of other farm inputs and environmental factors. Estimated coefficients indicate that a 10% increase in the genetic contribution of IRG accessions to an improved rice variety is associated with a yield increase of 27%. Through pedigree analysis, we also confirm that 45 to 77% of the genetic composition of improved rice varieties is derived from the genes of IRG accessions. Peta, Dee Geo Woo Gen, and Fortuna are the three most popular progenitors with definite IRG contribution. High genealogical diversity likely results from the crossing of germplasm received from multiple countries of origin, which also confer multiple, functional trait combinations in a released variety. Further, our calculations reveal that the average coefficient of parentage of all pairwise combinations among the 10 most adopted rice varieties is 0.0973, indicating a high degree of latent genetic diversity. Findings demonstrate the valuable contribution of the genetic resources conserved and distributed by IRG to the development of improved rice varieties.

Keywords Genebank · Genetic resources · Genetic diversity · Improved rice varieties · Productivity

1 Introduction

Rice (*Oryza sativa*) is the most important cereal crop and the staple food of more than half the world's population. Asia is the largest producing and consuming region (FAO 2014). During the Green Revolution of the 1960s, high-yielding rice varieties were introduced in response to the specter of famine as population densities rose and productivity stagnated. This remarkably successful process created an unintended consequence of crop diversity loss. Traditional varieties and wild species of rice were rapidly replaced by new varieties. Genetic variation from traditional varieties and related wild species is needed in crop improvement to cope with the many biotic and

abiotic stresses that continually challenge rice production around the world (IRRI 2018).

To protect against the loss of rice diversity, the International Rice Research Institute (IRRI) initiated the collection of rice genetic resources in 1962 and established the International Rice Genebank (IRG) in 1971. IRG has the largest and most diverse collection of rice genetic resources in the world. As of June 2018, the collection includes 130,139 accessions, comprising 123,837 *Oryza sativa*, 1655 of *O. glaberrima*, and 4647 accessions of wild relatives and interspecific hybrids. Since the 1970s, the rice genetic resources maintained by the IRG have been used to raise the productivity of rice cropping, particularly among smallholder farm families in lower income countries. Rice genetic improvement using IRG accessions is accomplished through IRRI's breeding programme. The breeding programme produces improved cultivars both in the form of varieties that are ready for immediate use by farmers and in the form of advanced lines suited for use as parent material in national plant-breeding programmes (Evenson and Gollin 1997, p. 471). However, other than the landmark studies by Evenson and Gollin (1997) and Gollin and Evenson (1998), we are not aware of analyses that have explicitly related farm productivity changes to genebank accessions through varietal improvement.

✉ Melinda Smale
msmale@msu.edu

Nelissa Jamora
nelissa.jamora@croptrust.org

¹ International Rice Research Institute (IRRI), Los Baños, Philippines

² Michigan State University, East Lansing, VA, USA

³ Global Crop Diversity Trust (Crop Trust), Bonn, Germany

Evenson and Gollin (1997) consulted the genealogies of 1709 rice varieties released by national programmes and IRRI from 1965 to 1990, correlating productivity changes with changes in IRRI programmes in an econometric model. They estimated that adding 1000 catalogued accessions was associated with the release of 5.8 additional varieties. Assuming a 10-year lag for variety development and a 10% discount rate, they calculated that these new accessions generated a present value (in 1990 dollars) of USD 325 million. High payoffs provided an economic justification for the continued operation of the International Network for the Genetic Evaluation of Rice and the IRG.

In another analysis, Gollin and Evenson (1998) examined impacts in India only. India is the world's second largest producer of rice, and more than half of India's population depends on rice as their staple food. The authors conducted a pedigree analysis of the 306 rice varieties released for planting in India over the period from 1965 to 1986. They applied a two-staged regression analysis to the district-level time series data to estimate the relative contribution of varietal improvement to productivity growth in rice. Their results showed that varietal change contributed more than one-third of the rice productivity gains realized over the post-Green Revolution period from 1972 to 1984. Their findings demonstrated that the economic value of genetic resources in India exceeds the costs of maintaining them (Gollin and Evenson 1998, p.149).

Two decades later, our analysis provides renewed, evidence-based documentation of the value of the IRG operation in raising productivity on farms in Eastern India. Unlike earlier studies, we are able to draw on data collected in a large-scale farm survey to test the genetic contribution of IRG accessions. We are also able to benefit from a digitized pedigree information system to better characterize ancestry and measure latent genetic diversity.

Our principal objective is to test the effect of IRG ancestry on the rice productivity of farmers in Eastern India in an econometric framework. We also examine the country of origin of all IRG accessions in the ancestry of improved rice varieties grown and characterize the most widely used progenitors. Finally, we evaluate the genealogical (latent) diversity of the most popular improved rice varieties grown by farmers using coefficients of parentage (COP). At the scale of a farming landscape, a negative relationship between productivity and crop diversity is often hypothesized (Widawsky and Rozelle 1998; Smale et al. 1998, 2008; Di Falco and Chavas 2009).

2 Data and methods

2.1 Data

Eastern India, comprising the states of Assam, Bihar, Jharkhand, Chhattisgarh, Orissa, eastern Uttar Pradesh, and

West Bengal, is the largest rice-growing region of the country. This region accounts for approximately 60% of the total rice area of India and generates around 48% of the national rice production (Adhya et al. 2008, p. 1). In 2016, IRRI conducted a farm-level survey in four states of Eastern India (eastern Uttar Pradesh, Bihar, Orissa, and West Bengal) through a project called the 'Rice Monitoring Survey' (RMS), funded by the Bill and Melinda Gates Foundation. This survey aimed to gather information on the rice varieties cultivated by farmers during the *kharif* or wet season of 2015 and examine diffusion of submergence as well as drought-tolerant rice varieties. A total of 720 villages were randomly selected from the rural villages defined in the 2011 Census of India and 12 households were randomly selected in each village. The total number of households interviewed was 8640 with the sample size in each state proportional to the rural population across states. Figure 1 shows the geographical location of the villages included in the survey. Table 1 shows the distribution of sample villages and household by state as well as the average rice area per household.

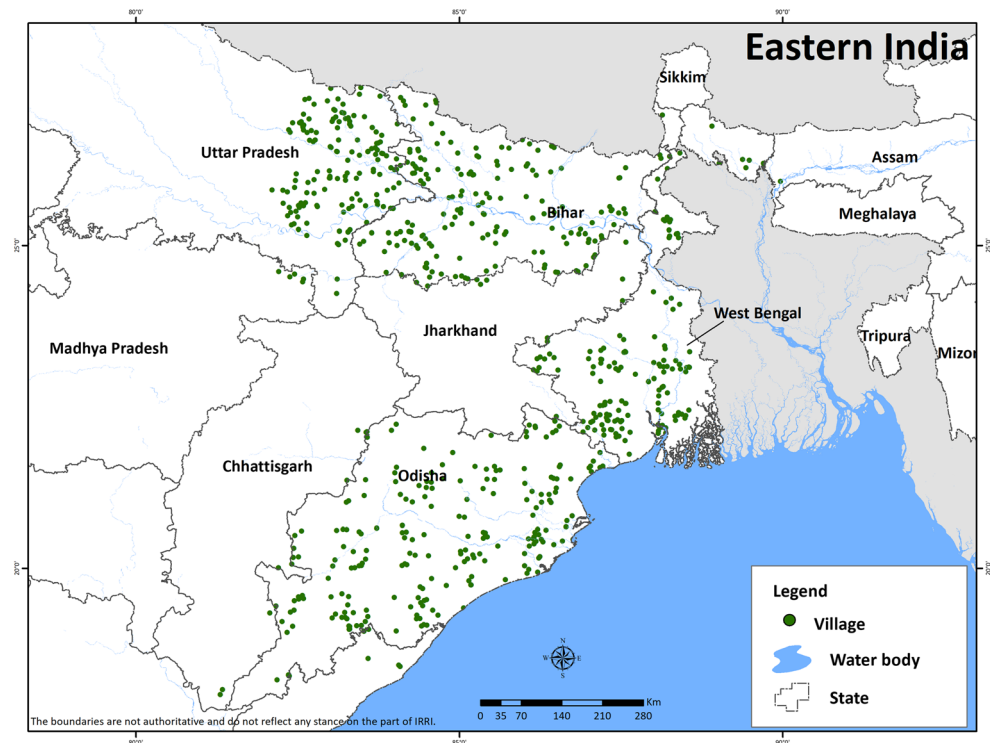
We used cross-sectional, farm-level data from RMS to identify the name of improved or released rice varieties grown by the farmers during the wet season of 2015 and constructed variables to explain productivity. Plot-level data was used in the econometric analysis, since there is one variety planted per plot and most of the farmers have more than one plot. However, only those plots that were planted with the improved rice varieties identified by this study and with pedigree information were included in the analysis. Those observations with outlier values based on extreme value analysis were also excluded from the sample. The final sample size used for model estimation was 8967 rice plots managed by 4298 farmers.

For the pedigree analysis of improved rice varieties, data were retrieved from existing databases, such as the International Rice Information System (IRIS), Genetic Resources Information Management System (GRIMS) and Genesys. IRIS is the rice information management system of the International Crop Information System (ICIS). ICIS provides integrated management of global information on genetic resources and crop cultivars and is used to manage germplasm information of materials developed, received, and maintained by IRRI. GRIMS is used to manage data from the different operations of the genebank, such as seed acquisition, multiplication, characterization, storage management, and seed distribution. Genesys is a gateway through which germplasm accessions from genebanks around the world can be easily found and ordered.

2.2 Methods

Descriptive analysis was conducted for all the improved rice varieties identified under the RMS project. Each improved rice variety was classified according to the breeding institution

Fig. 1 Geographical location of the sample villages of the RMS in 2016. Source: image depicted by authors and data from the 2016 RMS in Eastern India



responsible for its development and the source of its direct parent(s). Breeding institutions were classified as IRRI or non-IRRI. The source of direct parent(s) was classified as IRG or non-IRG. The improved rice varieties, which covered 95% of the total area planted to improved varieties, were identified using the results of the area estimates from RMS and became the focus of the succeeding analysis. Each of the identified varieties was classified according to whether it has direct parent(s) acquired from IRG.

Pedigree analysis was employed to quantify the genetic contribution of each IRG accession to each improved rice variety. Mendelgram, the programme of IRIS used for this analysis, assumes that each parent contributed an equal amount to their progeny. Using an algorithm that is consistent with Mendelian genetics, the genetic contribution is calculated as the probability that an unselected allele¹ comes from a progenitor with values ranging from 0 to 1.

In this study, the progenitor contribution of the IRG accessions was classified into four categories: definite contribution, possible contribution, no contribution, and unknown contribution. Definite contribution refers to identified progenitor contribution of the IRG accession in the ancestry of an improved rice variety. IRG accessions can be identified when the progenitor in the pedigree tree obtained in the BROWSE application of IRIS has an International Rice Germplasm Collection (IRGC) number opposite the variety name. This implies that the accession was obtained directly from IRG.

Possible contribution refers to a progenitor that does not have an IRGC number opposite the variety name in the pedigree tree, but it has a match by name somewhere in the IRG collection. In this case, it is possible that the progenitor came from IRG, but due to human error in encoding or a random error in the programme of the application/tool used in migrating and processing the data, the wrong name was recorded as the source. If a progenitor had no IRGC number opposite the variety name in the pedigree tree and had no match by name in the whole collection of the IRG, then it was classified as no contribution. Lastly, unknown contribution refers to the progenitor with unknown or confidential information.

Definite contribution was computed first during the aggregation of the progenitor contribution of IRG to each modern rice variety. In this process, the progenitors in the extracted pedigree tree with an IRGC number were identified and their ancestry levels were examined. If their pedigree lines were independent, then the progenitor contributions were added. However, in the case of recurrence of the IRGC accession in the same pedigree line, only the IRG accession with the lowest ancestry level was included in the aggregation to avoid double counting which could result in overestimation. Afterwards, following the same procedure for aggregating the progenitor contribution, the remaining progenitors in the extracted pedigree tree with names matched in the IRG collection were identified and their contributions were aggregated. This type of contribution is classified as the possible contribution. Similar procedures were employed for the aggregation of no and unknown contributions. The percent area of the improved

¹ *Allele* is a variant form of a gene.

Table 1 Distribution of sample households in Eastern India

State	Number of villages	Number of households	Average rice area (ha)
Eastern Uttar Pradesh	151	1812	0.61
Bihar	176	2112	0.81
Orissa	225	2699	0.74
West Bengal	168	2016	0.38
Total	720	8640	0.66

Source: data from the 2016 RMS in Eastern India

rice varieties adopted by farmers surveyed was used as a weight in computing the overall average progenitor contribution of IRG accessions by category.

A plot-level yield response function was estimated to test the effect of the genetic contribution of genebank accessions to productivity while controlling for the influence of conventional inputs (fertilizer, labor, machinery, plant protection, irrigation), management (age, education, access to inputs, credit and extension advice), and environmental factors (submergence, salinity, drought). The values of continuous variables pertaining to production output and input were transformed into unit per hectare. In the initial exploratory analysis, both definite and possible contributions of IRG on the variety were included in the model. However, in the final model, we only used the index for definite contribution to measure the clear impact of IRG on productivity.

Four functional forms were tested: linear, extended linear, Cobb-Douglas, and translog. Below are the specifications of these models:

1. Linear

$$y_i = \beta_0 + \sum_{i=1}^n \beta_i x_i + \sum_{i=1}^n \alpha_i z_i + \mu_i$$

2. Extended linear

$$y_i = \beta_0 + \sum_{i=1}^n \beta_i x_i + \sum_{i=1}^n \sum_{j=1}^n \beta_{ij} x_i x_j + \sum_{j=1}^n \alpha_j z_j + \mu_i$$

3. Cobb-Douglas

$$\ln y_i = \beta_0 + \sum_{i=1}^n \beta_i \ln x_i + \sum_{i=1}^n \sum_{j=1}^n \beta_{ij} \ln x_i \ln x_j + \sum_{j=1}^n \alpha_j z_j + \mu_i$$

4. Translog

$$\ln y_i = \beta_0 + \sum_{i=1}^n \beta_i \ln x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \beta_{ij} \ln x_i \ln x_j + \sum_{j=1}^n \alpha_j z_j + \mu_i$$

where y_i is yield of rice per hectare, x_i is the quantity of conventional production inputs per hectare, z_i is the non-conventional input variables (management, environmental

factors, and progenitor contribution), and μ_i is a random error.

Two different models were estimated per functional form. The first model did not control for location effects while the second model controlled for state effects. Initially, we tested a third model, which controlled for the village effects. However, due to small sample sizes per village, we excluded this model in the final analysis. Model diagnostics were performed to determine whether the necessary model assumptions are valid. We used Variance Inflation Factors (VIF) to test multicollinearity among independent variables and Breusch-Pagan/Cook-Weisberg test for heteroscedasticity. To evaluate the economic performance of the functional specification for each model, both Akaike's information criterion (AIC) and Bayesian information criterion (BIC) were applied. The most preferred model specification has the minimum AIC and BIC value. Stata 14.2 was used to estimate the model and to perform other necessary tests discussed above.

The key variable of interest in all models is the variable measuring genetic (progenitor) contribution. The null hypothesis is that the genetic contribution of IRG accessions does not affect rice yield. If the coefficient of this variable is significantly different from zero, then the null hypothesis is rejected and the alternative hypothesis that the contribution of IRG accession affects the yield of the improved rice varieties is accepted.

Similar econometric methodology has been employed in a small number of previous studies (e.g. Widawsky and Rozelle 1998; Smale et al. 1998, 2008; Di Falco and Chavas 2009). While these studies focused on the effects of variety diversity and genetic diversity on yield or yield risk, they did not link either outcome directly to genebank accessions through varietal improvement. To address this aspect, we also estimated econometric models using the preferred specification to test the effects of definite IRG contribution on the dependent variables of yield variance and yield kurtosis. Dependent variables were calculated from regression residuals of the main model, following Bozzola et al. (2018), which is based on Antle's method of moments (Antle 1987). The most popular IRG accessions in the ancestry of those adopted improved rice varieties were then identified. Initially, the progenitors with definite and possible contribution were segregated according to the decade when the variety was released. From this list, the study selected the three most common IRG accessions across decades. The

morphological characteristics of those IRG accessions were extracted from GRIMS and a table of varietal traits by IRG accession was constructed. These ‘profiles’ of variety characteristics supplement the econometric analysis.

Lastly, we selected for further genealogical analysis the 10 varieties most adopted by farmers during the 2015 season according to area planted. COP were computed for all pairwise combination among these 10 varieties to measure genetic diversity. The COP between two individuals is defined as the probability that a random allele at a random locus in one individual is identical by descent to a random allele at the same locus in the other individual (Cox et al. 1985, p. 529). The values of COP range between 0 (no known common ancestor) to 1 (same individual or variety). The lower the value of the COP, the higher the latent genetic diversity conferred by parentage among the included varieties (Souza et al. 1994, 1998). This final step provides evidence concerning the relationship between productivity and genetic diversity in the farming landscape during the survey year.

3 Results and discussion

3.1 Description of improved rice varieties grown by farmers

A total of 132 improved rice varieties (124 inbred and eight hybrid) were cultivated by rice farmers in Eastern India during the 2015 wet season. Five percent of these varieties were developed by IRRI, while the majority (80%) were developed by research institutions under the national breeding programme in India and other private companies. Non-IRRI-developed varieties covered around 78% of the total cultivated area of released varieties. Only 2% of area was planted to IRRI-developed varieties (for additional details, see Villanueva et al. 2019).

Of the 132 improved rice varieties cultivated by farmers during the 2015 wet season, 45 varieties covered 95% of the total cultivated area of improved rice varieties and are the focus of this analysis. Table 2 shows that 20% of these popular varieties have at least one direct parent that definitely came from the IRG. This comprises 43% of the total area. If possible IRG acquisition is included, which means that the direct parent has no IRGC number but has a name match with one of the IRG accessions, then 53% of the varieties have at least one direct IRG parent (corresponding to 71% of the total area).

3.2 Progenitor contribution of IRG

A total of 45 released varieties (inbred and hybrid) were cultivated on about 10.78 million ha. Fourteen of these varieties (six hybrid and eight inbred) do not have pedigree information

from the IRIS database, and thus, the progenitor contribution of IRG in these varieties is unknown.

Table 3 shows the summary of the progenitor contribution of IRG to the most adopted rice varieties by farmers during the 2015 wet season in Eastern India. Additional details per variety can be found in the Tables 10 and 11 in the Appendix. In definite, possible, and unknown progenitor contribution, the minimum and maximum contribution is 0 and 1, respectively, while the no contribution has a minimum of 0 and maximum of 0.50. The unweighted mean shows that the average definite and possible progenitor contribution of IRG to a released variety is about 35 and 25%, respectively. Considering the estimated percent planted area as weights, the average definite and possible progenitor contribution of IRG to a released variety is about 45 and 32%, respectively. These results mean that, on average, 45% of the genetic composition of a released rice varieties cultivated by farmers during the 2015 wet season came from IRG accessions. The progenitor contribution increases up to 77% if the possible contribution is added. This value can still be higher if there are IRG accessions in the pedigree of the hybrid and inbred varieties with no information. These results serve as evidence of the significant contribution of IRG to the development of improved rice varieties in Eastern India.

3.3 Country of origin

Table 4 shows the summary of the distribution of progenitors by country of origin and by type of progenitor contribution of IRG. Values shown in this table represent the aggregated frequency of unique progenitors of all released varieties by their country of origin. Here, progenitors that appeared in the pedigree of more than one progeny have been counted only once to avoid double counting. Results show that there are 122 unique progenitors from 18 countries with definite and possible contribution of IRG on released varieties in Eastern India cultivated during the 2015 wet season. Combining the results for definite and possible IRG contribution, the most popular country of origin is Philippines (35 progenitors), followed by India (30 progenitors) then the United States (15 progenitors) and Taiwan (8 progenitors). The majority of the progenitors from the Philippines were from IRRI where the IRG is located. The complete list of countries of origin of all ancestors by type of IRG contribution is shown in the Appendix.

These results demonstrate the diversity of the progenitors in terms of their country of origin, which can also translate into multiple trait combinations in a released variety. According to Ramirez et al. (2013, p. 44), the wide use of landraces from different countries as a source of desired traits has contributed to the increase in rice production in most rice-growing countries. The combined traits from these landraces conferred the necessary characteristics that allowed the different cultivars to cope with changing pest and disease pressures,

Table 2 Distribution of rice varieties adopted in Eastern India on 95% of area during the wet season of 2015 by IRG parentage

Parameters	Count	%	Area (000 ha)	%Area
With direct IRG parents (with IRGC# only)				
Both parents	2	4	124	1
One parent	7	16	4494	42
None	22	49	4485	42
Unknown	14	31	1677	16
Total	45	100	10,780	100
With direct IRG parents (with IRGC# and possible IRG acquisition)				
Both parents	8	18	623	6
One parent	16	36	7034	65
None	7	16	1447	13
Unknown	14	31	1677	16
Total	45	100	10,780	100

Source: data from the 2016 RMS in Eastern India and IRIS

various soil and nutrient conditions, and regional climatic conditions (Sebastian et al. 1998).

3.4 Contribution of IRG genetic resources to productivity

Table 5 reports the definition and basic descriptive statistics for the variables included in the yield response function. The average yield is around 2.5 t per hectare (t/ha) with standard deviation of 1.7 t/ha. Farmers applied an average of 153 kg/ha of nitrogen (N), phosphorus (P), and potassium (K) fertilizer on aggregate. Labor, power cost, and other material inputs has an average of 93 person-days/ha, Php5,521/ha, and Php778/ha, respectively. About half of the total plots experienced drought based on the perception of the farmer. The average age and education of the farmers is about 48 and six years, respectively. Only a few of the respondents have access to input, credit, and extension workers. The plot-level average of the definite contribution is 0.49, which means that about

49% of the genetic composition of the variety came from the genes of IRG accessions.

Villanueva et al. (2019) provide all estimated regression models and functional forms for each variable, including VIF. The range of VIF (1.01–3.47 with a mean of 1.55)

Table 4 Country of origin of progenitors with definite and possible IRG contribution during the 2015 wet season, Eastern India

Country of origin	Type of contribution		All
	Definite	Possible	
Philippines	13	22	35
<i>IRRI-Philippines</i>	8	20	28
<i>Philippine provinces</i>	5	2	7
India	12	18	30
United States	6	9	15
Chinese Taipei (Taiwan)	4	4	8
Indonesia	3	4	7
Malaysia	3	2	5
Thailand	4	3	7
Japan	0	2	2
Vietnam	2	1	3
Bangladesh	1	1	2
Ecuador	0	1	1
Korea, Republic Of	1	0	1
Pakistan	0	1	1
Gambia	1	0	1
China	0	1	1
Ivory Coast (Côte D'Ivoire)	0	1	1
Peru	1	0	1
Senegal	0	1	1
Total	51	71	122

Source: data from the 2016 RMS in Eastern India and IRIS

Table 3 Summary of the progenitor contribution of IRG to rice varieties adopted on 95% of area during the 2015 wet season, Eastern India

Type of contribution	Min	Max	Mean	
			Unweighted	Area-weighted
Definite	0.00	1.00	0.35	0.45
Possible	0.00	1.00	0.25	0.32
No	0.00	0.50	0.09	0.07
Unknown	0.00	1.00	0.31	0.16
Total			1.00	1.00

Source: data from the 2016 RMS in Eastern India and IRIS

Table 5 Definition of variables and some descriptive statistics

Variable name	Variable definition	Mean	SD
Yield	Yield (kg/ha)	2464	1718
Fertilizer	Fertilizer (NPK) quantity (kg/ha)	153	87
Labor	Total labor (person-days/ha)	93	55
Power_cost	Cost of power inputs (machine rental, animal rental, fuel and electricity) in Rs/ha	5521	5049
Other_inputs	Cost of other material inputs (herbicide, pesticide, and others) in Rs/ha	778	1370
Irrigation_percent	Percent of plot with irrigation	49	49
Crop_establishment	Crop establishment method (1-transplanting, 0-otherwise)	0.85	0.36
Submergence	Submergence condition of the plot (1-yes, 0-no)	0.05	0.23
Salinity	Salinity condition of the plot (1-yes, 0-no)	0.004	0.06
Drought	Drought condition of the plot (1-yes, 0-no)	0.52	0.50
Age	Age of the respondent (years)	48	13
Education	Education of the respondent (years)	6.0	4.5
Access_input	Access to material inputs (1-yes, 0-no)	0.12	0.32
Access_credit	Access to credit (1-yes, 0-no)	0.08	0.27
Access_extworker	Access to agricultural extension services (1-yes, 0-no)	0.07	0.26
Definite_contribution	Definite contribution of IRG on the variety (index)	0.49	0.35

SD = standard deviation. Rs/ha = Rupees per hectare

Source: data from the 2016 RMS in Eastern India and IRIS

confirms that there is a moderate multicollinearity among independent variables, but it is not severe enough to warrant corrective measure. We include other diagnostic tests for the four functional forms and each of the two models in Table 6. The result of the Breusch-Pagan/Cook-Weisberg test for heteroscedasticity is significant in each, which means that there is a problem of heteroscedasticity. To deal with this problem, robust standard errors were used in all models. Among all functional forms, results show that the translog has the lowest values of AIC and BIC. This result implies that translog is the best fit to the data. Between the two models with translog functional form, Model 2 was selected as the final model because it controls for state effects and captures significant yield variation across observations.

The preferred yield response model is presented in Table 7. The definite IRG contribution is the main variable of interest to us. The coefficient of the definite IRG contribution is positive and significant. This result means that the higher value of the genetic contribution from IRG in an improved rice variety increases its yield. The magnitude of the increase in yield can be computed by obtaining the exponential value of the coefficient of variable pertaining to the definite IRG contribution using the formula $\exp(\text{coefficient})$. After applying this formula, the result is 1.027. In terms of percent change, this result can be interpreted as a 1% increase in the definite IRG contribution on an improved rice variety can increase the yield by about 2.7%. Furthermore, a 10% increase in the definite IRG contribution can lead to a 27% yield increase. These findings

imply that the germplasm acquired from IRG is associated with yield improvement of rice varieties on farms in Eastern India.

Other results show that, except for labor, the conventional production inputs have positive and significant effects on productivity, conforming to the economic theory of production. The negative sign of labor input may be attributed to measurement error, as labor input was not adjusted with respect to the difference in quality between skilled and non-skilled farm workers. Person-days per hectare are often overstated, with variable lengths of workday. These findings were consistent when we explored different transformations of the labor variable and included them in the translog model. The sign of labor input remained negative. The signs of the other are consistent in various models that we evaluated.

Transplanted rice has significantly higher yield compared to broadcasted or direct seeded rice. Plots with a higher percentage of irrigated area have higher yield. Even in the wet season, some areas in Eastern India experience erratic rains and farmers need supplementary sources of irrigation for their crops. The coefficients of submergence, salinity, and drought are negative and significant. These abiotic stresses affect the growth of rice crops and decrease their productivity. In some parts of Eastern India, rice is cultivated in low-lying areas, which are prone to submergence due to heavy rains and intrusion of river or sea water. On the other hand, salinity can be caused by rising ground water or the intrusion of sea water, which brings salts to rice areas. Drought during the wet season

Table 6 Model diagnostic tests. Chi-squared value of Breusch-Pagan/Cook-Weisberg test for heteroscedasticity

Functional form	Model 1	Model 2
Linear	55.30 **	62.72 **
Extended linear	57.75 **	67.70 **
Cobb-Douglas	2,721.90 **	3,021.21 **
Translog	2,638.71 **	2,927.56 **

** significant at 1%

AIC and BIC

Functional form	AIC		BIC	
	Model 1	Model 2	Model 1	Model 2
Linear	154,670	154,283	154,784	154,418
Extended linear	154,424	154,073	154,623	154,293
Cobb-Douglas	38,609	38,358	38,723	38,493
Translog	38,472	38,202	38,670	38,422

Source: data from the 2016 RMS in Eastern India and IRIS

can be caused by untimely or low amounts of rainfall, and without other sources of water for irrigation, rice productivity will be affected.

Effects of management factors are weaker. The age and education of the farmers do not significantly affect the yield. Plots owned by farmers with access to credit have higher yield compared to those who do not have access. Credit can be in a form of farm input or cash that is used to provide the optimal needs in rice production. Finally, additional econometrical models testing the effects of definite IRG contribution on the yield variance and kurtosis (variability and downside risk) indicate risk-neutrality, which means that the contribution of progenitors does not significantly affect the yield risk.

3.5 Most popular progenitors

The progenitors that frequently appear in the pedigree of the released varieties in India during the 2015 wet season were enumerated by decade of varietal release. Peta, Dee Geo Woo Gen (DGWG), and Fortuna are the three most popular progenitors, with definite IRG contribution across the decades.

Peta and DGWG are the direct parents of IR8: the variety known as the ‘miracle rice’ that revolutionized rice production in tropical Asia. The development of IR8 was the start of the Green Revolution in rice not only because of rapid increase in rice production but also because this variety is the most widely used parent in several crosses in tropical Asia. On the other hand, Fortuna is one of the ancestors of elite varieties such as IR24, IR36, IR64, Lalat, Pooja and Swarna-Sub1. Table 8 ‘profiles’ the morphological and other special characteristics

Table 7 Estimation results, translog yield response function during the 2015 wet season, Eastern India

Independent variables	Coef.	Std. Err.	t
ln Fertilizer	0.869	0.21	4.15**
ln Labor	−0.673	0.31	−2.19*
ln Power_cost	0.139	0.08	1.78
ln Other_inputs	0.206	0.07	2.89**
(ln Fertilizer) ²	−0.034	0.01	−2.52*
(ln Labor) ²	0.236	0.03	7.80**
(ln Power_cost) ²	−0.002	0.00	−0.48
(ln Other_inputs) ²	0.015	0.00	3.20**
(ln Fertilizer) x (ln Labor)	−0.088	0.04	−2.32*
(ln Fertilizer) x (ln Power_cost)	0.027	0.01	3.06**
(ln Fertilizer) x (ln Other_inputs)	−0.040	0.01	−4.85**
(ln Labor) x (ln Power_cost)	−0.055	0.01	−3.87**
(ln Labor) x (ln Other_inputs)	−0.034	0.01	−3.65**
(ln Power_cost) x (ln Other_inputs)	0.013	0.00	4.21**
ln Irrigation_percent	0.127	0.01	10.41**
Crop_establishment	0.315	0.08	3.97**
Submergence	−2.469	0.16	−15.52**
Salinity	−1.279	0.51	−2.51*
Drought	−1.247	0.04	−29.66**
Age	−0.013	0.01	−1.35
Age x age	0.000	0.00	1.11
Education	−0.017	0.01	−1.15
Education * education	0.002	0.00	1.43
Access_input	−0.244	0.12	−2.04*
Access_credit	0.305	0.15	2.09*
Access_extworker	−0.128	0.12	−1.03
Definite_contribution	0.027	0.01	4.33**
I.state	Included		
Constant	4.764	1.02	4.66**

Source: data from the 2016 RMS in Eastern India and IRIS

No. of observations = 8967

Dependent variable = yield (kg/ha)

** significant at 1%, *significant at 5%

t = t-statistic, ln = natural logarithm, I.state = dummy variables for states in India

of these progenitors based on data assembled from the field and screen house evaluations. The following section provides additional details about Peta, DGWG, and Fortuna. These examples demonstrate that by conserving and distributing the seeds of key progenitors, IRG played a major role in producing widely adopted, improved rice varieties.

3.5.1 Peta (IRGC 35)

Peta, a tall and vigorous indica rice variety from Indonesia, was produced from a cross of Tjina and Latisail by H. Siregar. Tjina is presumed to originate from China, while Latisail came

Table 8 Morphological characteristics of popular IRG progenitors of improved rice varieties during the 2015 wet season, Eastern India

Characteristics	IRG accessions		
	Peta	DGWG	Fortuna
Variety group	Indica	Indica	Javanica
Country of origin	Indonesia	Taiwan	Taiwan
Date of acquisition	–	24-Mar-1962	7-Jul-1978
Maturity (days)	145	111	124
Seedling height	Intermediate (30–59 cm)	Short (<30 cm)	Intermediate (30–59 cm)
Grain length (mm)	9.3	8.1	9.8
Grain width (mm)	2.9	3.1	3.0
100-grain weight (gm) – cultivated	2.9	2.3	3.1
Leaf length – cultivated	Intermediate (41–60 cm)	Intermediate (41–60 cm)	Intermediate (41–60 cm)
Leaf width – cultivated	Intermediate (1–2 cm)	Intermediate (1–2 cm)	Intermediate (1–2 cm)
Panicle length (cm) at post-harvest	28.6	24	32
Panicle type	Semi-compact	Drooping	Spreading (open)
Endosperm type	–	Non-glutinous	Non-glutinous
Response on diseases			
Bacterial blight	Susceptible	–	Susceptible
Blast	Moderately resistant	Moderately resistant	Susceptible
Sheath blight	Moderately susceptible	–	–
Rice tungro virus (% infected)	34	92	89
Rice ragged stunt virus (% infected)	61	89	71
Response on insects			
Brown planthopper	Susceptible	Susceptible	Susceptible
Green leafhopper	Intermediate	Susceptible	Susceptible
Rice whorl maggot	Susceptible	Susceptible	Susceptible
White-backed planthopper	Susceptible	Susceptible	Susceptible
Zigzag leafhopper	Resistant	–	Susceptible
Rice leaf folder	Susceptible	–	Susceptible
Yellow stemborer	Susceptible	–	Susceptible
Striped stemborer	–	Susceptible	–
Drought tolerance			
Early vegetative stage	–	Intermediate	Resistant
Late vegetative stage	Intermediate	–	Intermediate
Reproductive stage	Susceptible	Susceptible	–
Flood tolerance	Susceptible	–	Susceptible
Salt tolerance	Tolerant	–	Intermediate
Cold tolerance	Susceptible	Susceptible	Susceptible

Source: data from the 2016 RMS in Eastern India and IRIS

from the Bengal region (Dalrymple 1986). The IRG first acquired Peta in the early 1960s, although the exact date was not documented. IRG assigned an IRGC number of 35 for this acquisition. Other Peta seed samples were acquired by IRG in 1961, 1971, 1972, and 1976. We focus on IRGC 35 since it is the accession IRRI breeders used to produce IR8.

Peta is a tall (28.6 cm) and late-maturing variety (145 days) with a semi-compact panicle at post-harvest stage. It has an intermediate size of leaf length and width. Its grain length

and width are 9.3 and 2.9 mm, respectively, with grain weight per 100 grains of 2.9 g. Based on the results of field evaluation on biotic stresses, Peta is resistant to tungro virus and moderately resistant to blast. This variety is susceptible to bacterial blight and moderately susceptible to sheath blight and ragged stunt virus. Peta is also susceptible to all destructive insects, except the zigzag leafhopper. For abiotic stresses, it is tolerant to saline conditions but susceptible to drought, floods, and cold.

We searched the elite lines and varieties where Peta was one of the ancestors from first up to fourth degree of their pedigrees. We found 5728 unique advanced lines and released varieties with Peta ancestry in IRIS, most of which were developed in the 1960s. The high frequency of Peta in their ancestries is due to the crossing of IR8 with other varieties or landraces to produce more improved rice varieties. Aside from IR8, other most notable released varieties with Peta in their ancestries are IR36, IR42, IR64, IR72, Swarna, Pooja, and Lalat. Twenty out of 31 popular varieties (65%) cultivated during the 2015 wet season in Eastern India and with pedigree data have Peta as one of their ancestors. These varieties were planted in around 6.6 million ha during that season, representing 58% of the total area of improved rice varieties in Eastern India.

3.5.2 DGWG (IRGC 123)

DGWG, found in Taiwan, is the earliest known semi-dwarf rice and is also known as 'I-geo-woo-gen'. The prefixes Dee-geo and I-geo mean 'dwarf' (Dalrymple 1986, p. 17). The origins of DGWG are unclear. One account suggests that it may have been brought from Fujian several hundred years ago (Miu 1959, p. 67), while another suggests that it may have been a spontaneous mutant from another traditional variety named Woo-gen (Hu 1973 p. 566). DGWG soon became popular in Taiwan and was planted on 10,907 ha during the first cropping season in 1953.

Before the development of IR8 in early 1960s, DGWG was used in the first cross involving a semi-dwarf in Taiwan, the 'Tsai-yuan-chung': a tall, disease-resistant local variety. A selection from this cross was named 'Taichung Native 1' (TN-1) in 1956. TN-1 was rapidly accepted by farmers due to its short-stature and high tillering. By 1965, 79,000 ha were planted with TN-1, making it the second most popular variety that year (Dalrymple 1986, p. 18).

On 24 March 1962, IRG acquired DGWG seeds from Taiwan and assigned 123 as its IRGC number. Afterwards, this variety became the donor of the dwarfism trait in IR8. Dr. Robert Chandler, former director general of IRRI, described DGWG as 'a high-yielding, heavy-tillering, short-statured variety from Taiwan' (Hargrove and Coffman 2006, p. 36).

DGWG is a medium maturing (111 days) variety. It has intermediate size of leaf in terms of length and width. Its grain has a length of 8.1 mm, width of 3.1 mm, and 100-grain weight of about 2.3 g. It has a short and drooping panicle with a length of 24 cm. DGWG was categorized as non-glutinous or non-waxy rice based on the starch in the endosperm. It is moderately resistant to blast but very susceptible to tungro and ragged stunt virus; drought and cold conditions; and pests, such as the brown planthopper, green leafhopper, whorl maggot, white-backed planthopper, and striped stemborer.

According to the database, DGWG with IRGC 123 is present in the first up to fourth pedigree of 5649 advanced lines and released varieties, which shows its immediate contribution in developing improved rice varieties. Some of the notable released varieties are IR6, IR8, CO 36, and Giza 180. Similar to Peta, the popularity of DGWG is due to the development of IR8, which was crossed with other varieties to produce new improved rice varieties. Some of the well-known released varieties that are progenies of DGWG are IR36, IR42, IR64, IR72, Swarna, Pooja, and Lalat. DGWG was found in the ancestry of 15 out of 31 popular varieties (48%) with pedigree information. These varieties were cultivated on around 6.2 million ha during the 2015 wet season in Eastern India, covering 55% of the total area of improved rice varieties.

3.5.3 Fortuna (IRGC 139)

Fortuna is a landrace javanica rice that originated in Taiwan. The IRG acquired it on 7 July 1978 and assigned it an IRGC number of 139 (DOI <https://doi.org/10.18730/1PME6>). Fortuna is a late-maturing variety (124 days) with intermediate plant height. It has a grain length of 9.8 mm, width of 3.0 mm, and a 100-grain weight of 3.1 g, which is similar to Peta. The variety has an intermediate size of leaf with a long (32 cm) and spreading panicle (open). Based on its endosperm, Fortuna is classified as non-glutinous or non-waxy rice. This variety is susceptible to bacterial blight, sheath blight, tungro virus, and ragged stunt virus as well as all common destructive insects. In the early vegetative stage, it is resistant to drought, but it becomes susceptible in the late vegetative stage. While it exhibits intermediate tolerance to salinity condition, it is susceptible to flooding and cold conditions.

Fortuna with IRGC 139 was found in the first up to fourth degree of the pedigree of 765 advanced lines and released varieties, illustrating its immediate genetic contribution to improved rice varieties. Some of the noteworthy varieties in this list are Blue Bonnet, Sigadis, Star bonnet, and Sun bonnet. These varieties were then crossed with other improved rice varieties and elite lines to develop other improved rice varieties, including the popular varieties Milfor, IR64, Lebonnet, IR72, Lalat, Pooja, and Swarna-Sub1. Among the 31 popular improved rice varieties with pedigree information, we found that 11 varieties (35%) have Fortuna in their ancestries. During the 2015 wet season, these varieties were cultivated on around 2.5 million ha in Eastern India, amounting to 22% of the total area of improved rice varieties.

3.6 Latent genetic diversity

Table 9 shows the off-diagonal matrix of the COP of the pairwise combinations among the 10 most adopted varieties by farmers

Table 9 Coefficient of parentage of 10 most adopted varieties with pedigree information during the 2015 wet season, Eastern India

Variety name	Swarna	Mahsuri	Pooja	Lalat	Sarjoo52	Vijetha (MTU 1001)	Moti	Sambha Mahsuri	Swarna-Sub1	Cottondora Sannalu (MTU 1010)
Swarna	1.0000									
Mahsuri	0.3125	1.0000								
Pooja	0.0547	0.0000	1.0000							
Lalat	0.0513	0.0000	0.0521	1.0000						
Sarjoo 52	0.0313	0.0000	0.0391	0.0330	1.0000					
Vijetha (MTU 1001)	0.0833	0.1250	0.0934	0.0555	0.0869	1.0000				
Moti	0.0039	0.0000	0.0076	0.0050	0.0313	0.0119	1.0000			
Sambha Mahsuri	0.1719	0.5000	0.0156	0.0203	0.1250	0.1068	0.0156	1.0000		
Swarna-Sub1	0.9415	0.2930	0.0556	0.0546	0.0320	0.0849	0.0041	0.1630	1.0000	
Cottondora Sannalu (MTU 1010)	0.0915	0.1250	0.1036	0.0558	0.0358	0.1207	0.0069	0.0852	0.0926	1.0000

Source: data from the 2016 RMS in Eastern India and IRIS

with pedigree information during the 2015 wet season in Eastern India. All values in the diagonal of the half matrix are 1 since the progenitors between a variety and itself are perfectly the same. Among the 45 pairwise combinations of these varieties, 34 of them (76%) have values of COP less than or equal to 0.10. Four varieties (Pooja, Lalat, Sarjoo 52, and Moti) have no identical progenitors with Mahsuri because the values of their COP are zero. On the other hand, the pairwise combination with the highest value of COP is between Swarna and Swarna-Sub1 (0.94) followed by the combination of Mahsuri and Sambha Mahsuri (0.50). Swarna-Sub1 is closely related to Swarna because Swarna-Sub1 is the improved version of Swarna after adding the *Sub1* gene, which makes this variety tolerant to submergence. Samba Mahsuri is the progeny after crossing two Indian popular varieties, Sambha and Mahuri, which is why 50% of their progenitors are identical.

On average, the COP for all pairwise combinations among these varieties (excluding the COP of a variety to itself) is 0.0973. This means that two varieties have a mean of 9.73% identical progenitors in their pedigrees, suggesting a high degree of latent genetic diversity. High diversity among these varieties is likely the result of crossing germplasm from different countries of origin, as shown in the previous section. This ancestral diversity may also be reflected in diverse, multiple trait combinations that provide functional diversity. High varietal diversity may reduce yield variability when pest infestations strike or bad weather occurs (Widawsky and Rozelle 1998).

4 Conclusion

We have provided new evidence concerning the contribution of IRG to rice productivity among smallholder farms in Eastern India, following earlier studies by Evenson and Gollin (1997) and Gollin and Evenson (1998). The unique feature of this study is that we were able to combine data collected in personal farmer interviews with digitized pedigree data to test the role of IRG ancestry in farm productivity in

Eastern India. To accomplish this, we identified the nature and origins of the IRG ancestry in varieties covering 95% of the rice area surveyed. We accessed other characterization data from field and screenhouse evaluations to ‘profile’ the three most popular IRG progenitors with definite IRG contribution over the decades. Lastly, we computed the COP for all pairwise combinations among the 10 improved rice varieties most adopted by farmers.

We found that, on average, 45% of the genetic composition of an improved rice variety cultivated by farmers in Eastern India was definitely contributed by IRG accessions. If possible IRG contributions are included, the total progenitor contribution could increase up to 77%. To assess the farm-level impact, the index for the definite IRG contribution was included as one of the explanatory variables of the yield response function. The results of the translog model show that the definite IRG contribution has a positive and significant impact on yield. A 1% increase in the definite IRG contribution to an improved rice variety increases rice yield by about 2.7%, other factors held constant.

Based on the country of origin of the progenitors, there are 122 unique progenitors from 19 countries with definite and possible IRG contribution. This demonstrates the diversity of the IRG progenitors in terms of their country of origin, implying that the wide use of combined traits from the landraces originating in different countries may have conferred the necessary characteristics that allowed farmers to cope with biotic and abiotic conditions and raise their rice production.

The most popular IRG progenitors with definite contribution identified in this study are Peta, DGWG, and Fortuna. Peta and DGWG were popular progenitors because they are the direct parents of IR8 and the development of IR8 was the start of the Green Revolution on rice. Twenty out of 31 popular varieties (65%) identified by this study have Peta, DGWG or Fortuna in their ancestry. These 20 varieties were planted on around 6.6 million ha during the 2015 wet season, covering 58% of the total area of improved rice varieties in Eastern India.

At the same time, COP analysis indicates that the top 10 most adopted varieties have only 9.73% identical progenitors in their

pedigrees, implying a high degree of diversity conferred by ancestry. High diversity among these varieties probably reflects the crossing of germplasm sources from different countries of origin and may also imply functionally diverse, multiple trait combinations. Consistent with this result, we find that the definite IRG contribution is neutral to yield risk in rice production.

The findings of this study demonstrate the valuable contribution of IRG's conservation and distribution of genetic accessions to the development of improved rice varieties and rice production on farms in Eastern India. Future research might include applications to other farming areas, preferably with panel data to enhance the robustness of the analysis.

Acknowledgements Funding for this research was provided by the CGIAR Genebank Platform, IRRI, and the Crop Trust through the

2018 Genebank Impacts Fellowship. We would like to acknowledge the IRG and the RMS team for sharing the data.

Authorship contribution The first author contributed to the research conceptualization and design, data gathering, data analysis, writing, and editing. The fourth and fifth authors contributed to the research conceptualization and design and data provision. The corresponding authors contributed to the research conceptualization and design, data analysis, writing, and editing.

Compliance with ethical standards

Conflict of interest The first author is currently an economist at IRRI. The third author is an agricultural economist at the Crop Trust. The fourth author is currently a staff member at the IRG. The fifth author was the Genebank Manager of IRG at the time this research was conducted. The second author declares no conflict of interest.

Appendix

Table 10 Varietal information and the progenitor contribution of IRG to rice varieties adopted on 95% of area during the 2015 wet season, Eastern India

Variety name	Released Year	Variety type	Cultivated area (000 ha)	% Area	Progenitor contribution			
					Definite	Possible	No	Unknown
Swarna	1979	Modem/HYV	3592.58	31.68	0.7500	0.2500	0.0000	0.0000
Mahsuri	1972	Modem/HYV	1254.79	11.07	0.0000	1.0000	0.0000	0.0000
Arize 6444	2001	Hybrid	840.56	7.41	0.0000	0.0000	0.0000	1.0000
Sarjoo 52	1980	Modem/HYV	436.31	3.85	0.0000	0.6250	0.3750	0.0000
Pooja	1999	Modem/HYV	712.84	6.29	0.8750	0.1250	0.0000	0.0000
Lalat	1988	Modem/HYV	572.86	5.05	0.6172	0.0078	0.3750	0.0000
Moti	1988	Modem/HYV	222.08	1.96	0.2813	0.4063	0.3125	0.0000
Moti gold	2010	Modem/HYV	201.24	1.77	0.0000	0.0000	0.0000	1.0000
Samba Mahsuri	1986	Modem/HYV	168.71	1.49	0.2500	0.7500	0.0000	0.0000
Vijetha (MTU 1001)	1995	Modem/HYV	421.06	3.71	0.3936	0.3493	0.2572	0.0000
Sonam		Modem/HYV	131.91	1.16	0.0000	0.0000	0.0000	1.0000
CR Dhan 501	2010	Modem/HYV	79.13	0.70	0.5000	0.5000	0.0000	0.0000
Swarna- Sub1	2009	Modem/HYV	158.81	1.40	0.7452	0.2548	0.0000	0.0000
GK 5003	2008	Hybrid	90.15	0.79	0.0000	0.0000	0.0000	1.0000
Cottondora Sannalu (MTU 1010)	2000	Modem/HYV	157.34	1.39	0.5165	0.1671	0.3164	0.0000
Pratikshya	2005	Modem/HYV	151.55	1.34	0.8684	0.1277	0.0039	0.0000
Samba-Sub1	2003	Modem/HYV	50.44	0.44	0.0000	0.0000	0.0000	1.0000
IR64-Sub1	2014	Modem/HYV	137.37	1.21	0.8305	0.0533	0.1162	0.0000
Niranjan		Modem/HYV	125.76	1.11	0.0000	0.0000	0.0000	1.0000
Sita	1972	Modem/HYV	70.85	0.62	1.0000	0.0000	0.0000	0.0000
Annapoorna	1976	Modem/HYV	74.52	0.66	0.5000	0.0000	0.5000	0.0000
Jaya	1968	Modem/HYV	112.48	0.99	0.0000	1.0000	0.0000	0.0000
Khandagiri	1992	Modem/HYV	86.38	0.76	0.6152	0.3848	0.0000	0.0000
Ranjeet	1994	Modem/HYV	90.65	0.80	1.0000	0.0000	0.0000	0.0000
Geetanjali	2005	Modem/HYV	87.80	0.77	0.0000	1.0000	0.0000	0.0000
Naveen	2005	Modem/HYV	77.67	0.68	0.1250	0.6250	0.2500	0.0000

Table 10 (continued)

Variety name	Released Year	Variety type	Cultivated area (000 ha)	% Area	Progenitor contribution			
					Definite	Possible	No	Unknown
Udayagiri	1999	Modern/HYV	74.29	0.66	0.4922	0.5078	0.0000	0.0000
Pioneer		Hybrid	44.56	0.39	0.0000	0.0000	0.0000	1.0000
Mehak		Modern/HYV	22.43	0.20	0.0000	0.0000	0.0000	1.0000
Bhuban	1988	Modern/HYV	53.03	0.47	0.5000	0.5000	0.0000	0.0000
Krishna	1970	Modern/HYV	31.80	0.28	0.5000	0.5000	0.0000	0.0000
Super Shyamali		Modern/HYV	55.84	0.49	0.0000	0.0000	0.0000	1.0000
Tulasi	1988	Modern/HYV	49.04	0.43	0.1250	0.3750	0.5000	0.0000
Jamuna	1970	Modern/HYV	44.22	0.39	0.0000	1.0000	0.0000	0.0000
Miniket		Modern/HYV	33.45	0.29	0.0000	0.0000	0.0000	1.0000
Santosh	2001	Modern/HYV	41.58	0.37	0.5000	0.0000	0.5000	0.0000
Gorakhnath 509	2011	Hybrid	14.50	0.13	0.0000	0.0000	0.0000	1.0000
IR 36	1981	Modern/HYV	28.30	0.25	0.9688	0.0313	0.0000	0.0000
Arzel		Modern/HYV	30.49	0.27	0.0000	0.0000	0.0000	1.0000
Pankaj	1969	Modern/HYV	33.40	0.29	1.0000	0.0000	0.0000	0.0000
Kaveri 108 (Sampurna)		Hybrid	10.69	0.09	0.0000	0.0000	0.0000	1.0000
NK 5251	2009	Hybrid	24.53	0.22	0.0000	0.0000	0.0000	1.0000
Savithri	1983	Modern/HYV	28.03	0.25	0.5000	0.5000	0.0000	0.0000
Annada	1987	Modern/HYV	29.62	0.26	0.5000	0.0000	0.5000	0.0000
Sahbhagi Dhan	2009	Modern/HYV	24.38	0.21	0.7809	0.1937	0.0254	0.0000

HYV = high-yielding variety. Source: data from the 2016 RMS in Eastern India and IRIS

Table 11 Most popular progenitors of popular improved varieties with definite and possible IRG contribution by decade of varietal release during the 2015 wet season, Eastern India

Year of release	Type of progenitor contribution		All
	Definite	Possible	
1968–1970	GEB 24, Peta, Tangkai Rotan	DGWG, TN-1, Cina, Latisail, T 141, Basmati 370	DGWG, TN-1, Cina, Latisail, T 141, Basmati 370, GEB 24, Peta, Tangkai Rotan
1971–1980	DGWG, Peta, I Geo Tze, IR 12–178–2-3, Mahsuri, Mong Chim Vang A	Cina, IR 8, Latisail, Kameji, Mayang Ebos 80, Shinkiri, Taichung 65	DGWG, Cina, IR 8, Latisail, Peta
1981–1990	Peta, DGWG, TKM 6, B 5580 A 1–15, Benong, Bluebonnet, Century Patna 231, CO 13, CR 94–13, Fortuna, IR 579–48–1-2 (Nilo 11), Sigadis, Tadukan, W 1543	TN-1, DGWG, Blue Rose Supreme, Cina, Latisail, Rexoro, Sinampaga, Texas Patna	TN-1, DGWG, Peta, Blue Rose Supreme, Cina, Latisail, Rexoro, Sinampaga, Texas Patna, Century Patna 231, TKM 6
1991–2000	Peta, DGWG, Fortuna, B 5580 A 1–15, Benong, Bluebonnet, Century Patna 231, Gam Pai 30–12-15, IR 579–48–1-2 (Nilo 11), Sigadis, Tadukan, TN-1, TKM 6, W 1543	Cina, Latisail, IR 8, Rexoro, Blue Rose Supreme, CO 18, Gam Pai, IR 24, IR 262–43–8-11, Sinampaga, SLO 17, Texas Patna	Cina, Latisail, Fortuna, Peta, IR 8, Rexoro, DGWG
2001–2014	Peta, B 5580 A 1–15, Benong, Bluebonnet, Century Patna 231, DGWG, Fortuna, Seraup Besar 15, Sigadis, W 1543	Cina, Latisail, Fortuna, Blue Rose Supreme, CO 18, IR 24, IR 8, Rexoro, Sinampaga, SLO 17, Texas Patna	Fortuna, Cina, Latisail, Peta, DGWG
All	Peta, DGWG, Fortuna, B 5580 A 1–15, Benong, Bluebonnet, Century Patna 231, Sigadis, TKM 6, W 1543	Cina, Latisail, IR 8, Rexoro, TN-1, Blue Rose Supreme, Sinampaga, Texas Patna, CO 18, IR 24, SLO 17	DGWG, Peta, Cina, Latisail, TN-1, Fortuna, IR 8, Rexoro

Source: data from the 2016 RMS in Eastern India and IRIS

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adhya, T. K., Singh, O. N., Swain, P., & Ghosh, A. (2008). Rice in eastern India: Causes for low productivity and available options. *Journal of Rice Research*, 2, 1–5.
- Antle, J. M. (1987). Econometric estimation of producers' risk attitudes. *American Journal of Agricultural Economics*, 69, 509–522.
- Bozzola, M., Smale, M., & Di Falco, S. (2018). Climate shocks, weather, and maize intensification decisions in rural Kenya. In C. S. Berck, P. Berck, & S. Di Falco (Eds.), *Agricultural adaptation to climate change in Africa: Food security in a changing environment* (pp. 107–128). New York: RFF Press/Routledge.
- Cox, T. S., Kiang, Y. T., Gorman, M. B., & Rodgers, D. M. (1985). Relationship between coefficient of parentage and genetic similarity indices in the soybean. *Crop Science*, 25, 529–532.
- Dalrymple, D. (1986). Development and spread of high-yielding rice varieties in developing countries. *Bureau for Science and Technology*, 33, 17–18.
- Di Falco, S., & Chavas, J. P. (2009). On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia. *American Journal of Agricultural Economics*, 91(3), 599–611.
- Evenson, R. E., & Gollin, D. (1997). Genetic resources, international organisations, and improvement in rice varieties. *Economic Development and Cultural Change*, 45(3), 471–500.
- FAO. (2014). *FAO statistical yearbook 2014, Asia and the Pacific food and agriculture*. Bangkok: Food and Agriculture Organization of the United Nations Regional Office for Asia and the Pacific.
- Gollin, D., & Evenson, R. E. (1998). An application of hedonic pricing methods to value rice genetic resources in India. In R. E. Evenson, D. Gollin, & V. Santaniello (Eds.), *Agricultural values of plant genetic resource* (pp. 139–150). Wallingford: CAB International.
- Hargrove, T., & Coffman, W. R. (2006). Breeding history. *Rice Today*, 5(4), 35–38.
- Hu, C. H. (1973). Evaluation of breeding semi-dwarf rice by induced mutation and hybridization. *Euphytica*, 22(3), 562–574.
- International Rice Research Institute (IRRI). (2018). The International Rice Genebank. <http://irri.org/our-work/research/genetic-diversity/international-rice-genebank>.
- Miu, T. S. (1959). A photographic monograph of rice varieties of Taiwan. *Taiwan Agricultural Research Institute, Special Publication*, 2, 67.
- Ramirez, M., Ortiz, R., Taba, S., Sebastian, L., Peralta, E., Williams, D., et al. (2013). Demonstrating interdependence on plant genetic resources for food and agriculture. In M. Halewood, N. I. Lopez, & S. Louafi (Eds.), *Crop genetic resources as a global commons: Challenges in international law and governance* (pp. 39–61). London: Routledge.
- Sebastian, L. S., Hipolito, L. R., & Garcia, J. (1998). Molecular contribution of landraces in the development of Philippine-bred modern rice varieties. *SABRAO Journal of Plant Breeding and Genetics*, 30(2), 73–82.
- Smale, M., Hartell, J., Heisey, P., & Senauer, B. (1998). The contribution of genetic resources and diversity to wheat productivity and stability in the Punjab of Pakistan. *American Journal of Agricultural Economics*, 80, 482–493.
- Smale, M., Singh, J., Di Falco, S., & Zambrano, P. (2008). Wheat breeding, productivity and slow variety change: Evidence from the Punjab of India after the green revolution. *Australian Journal of Agricultural and Resource Economics*, 6(52), 419–432.
- Souza, E., Fox, P. N., Byerlee, D., & Skovmand, B. (1994). Spring wheat diversity in irrigated areas of two developing countries. *Crop Science*, 34(3), 774–783.
- Souza, E., Fox, P. N., & Skovmand, B. (1998). Parentage analysis of international spring wheat yield nurseries 17 to 27. *Crop Science*, 38, 337–341. <https://doi.org/10.2135/cropsci1998.001183X003800020009x>.
- Villanueva, D., Capilit, G. L., Smale, M., Jamora N., & Sackville-Hamilton, R. (2019). The contribution of the international Rice Genebank (IRG) to varietal improvement and farm productivity in eastern India. *Genebank Impacts Working Paper No. 3*. CGIAR Genebank platform, IRRI, and the crop trust.
- Widawsky, D., & Rozelle, S. (1998). Varietal diversity and yield variability in Chinese rice production. In M. Smale (Ed.), *Farmers, gene banks, and crop breeding* (pp. 159–172). Boston: Kluwer.



Donald Villanueva is currently an Associate Scientist working on impact assessment projects in the Impact Evaluation and Learning (IEL) cluster under Agri-food Policy Platform in the IRRI. While working at IRRI, he was selected as one of the Genebank Impacts Fellows of the Crop Trust in 2018 to evaluate the impact of the IRG on varietal improvement in India. He has an MSc in Economics and BSc in Statistics from the University of the Philippines Los Baños.

Before joining the IEL cluster in 2017, he served as an Assistant Scientist under the Gender Research Team of IRRI where he was involved in the data collection, validation, and analysis of gender-related studies in South and Southeast Asia.



Melinda Smale joined the Food Security Group at Michigan State University (MSU) in 2011 as Professor of International Development, working intensively with research partners in Kenya, Zambia and Mali. From 2002, at the International Food Policy Research Institute (IFPRI) and Bioversity International, she led a global research programme about the economics of crop genetic resources. The team studied biotech crops, genebanks, on-farm crop diversity, local seed

markets, and underutilized crops in numerous countries, including Mali, Uganda, India, Nepal, Bolivia, Colombia and the Philippines. While living in Malawi and later in Mexico (1989–2000), she analyzed the adoption and impacts of improved wheat and maize seed for the International Maize and Wheat Improvement Center (CIMMYT). During the 1980s,

Melinda worked in Pakistan, Somalia, Mauritania and Niger on shorter-term assignments. She has published over 100 peer-reviewed articles, has received awards for outstanding articles and has served on the editorial committees of several journals. She enjoys working with young professionals.



Nelissa Jamora leads the monitoring, evaluation, and impact assessment of programmes and projects of the Crop Trust. She has a PhD in Agricultural Economics from the University of Göttingen in Germany and an MSc from Michigan State University in the USA. Her interest in agricultural development was motivated from her early work at IRRI on price and market analysis and impact assessment involving rice farmers in the Philippines. Nelissa received a USA Fulbright scholarship

in 2004 and a grant for early-researchers in 2014 from the German Research Foundation. She was recognized as a Young Rice Scientist at the International Rice Congress in 2014 for her work on rice price analysis. She joined the Crop Trust in 2015 and currently coordinates the Community of Practice on Genebank Impacts.



Grace Lee Capilit is a Lead Specialist in data management at IRRI. She currently manages the rice germplasm database of the International Rice Genebank and collaborates with scientists and researchers on genebank operations, such as seed acquisition, multiplication, characterization, conservation, seed distribution, and information sharing. She has an MS and a BS degree in Statistics from the University of the Philippines Los Baños (UPLB). Before joining the TT

Chang Genetic Resources Center in 2009, she was a database manager at the Plant Breeding, Genetics and Biotechnology Division of IRRI from 2000 to 2008 and also served as a teaching associate at UPLB handling statistical programming, experimental designs, use of statistical packages, and general statistics classes.



Ruaraidh Sackville Hamilton has over 40 years of experience in the conservation and use of crop genetic resources, including best practices and workflow management systems for genebank management; database design and data management; statistics, genetics, and genomics; crop wild relatives; pre-breeding; plant breeding; plant ecology; GM biosafety; and international policy on access and benefit-sharing. From 2002 to 2018, he served as the Head of the International Rice

Genebank at IRRI in the Philippines. He is an evolutionary biologist and one of the seven award recipients of the Crop Trust Legacy Award in 2018 for his effort and outstanding contribution in the field of plant genetic resources conservation. He has a PhD in plant genetic resources from the University of Cambridge, UK.