

# Genetic Trends Estimation in IRRIs Rice Drought Breeding Program and Identification of High Yielding Drought-Tolerant Lines

**Apurva Khanna**

IRRI: International Rice Research Institute

**Mahender Anumalla**

IRRI: International Rice Research Institute

**Margaret Catolos**

IRRI: International Rice Research Institute

**Jérôme Bartholomé**

IRRI: International Rice Research Institute

**Roberto Fritsche-Neto**

IRRI: International Rice Research Institute

**John Damien Platten**

IRRI: International Rice Research Institute

**Daniel Joseph Pisano**

IRRI: International Rice Research Institute

**Alaine Gulles**

IRRI: International Rice Research Institute

**Ma Teresa Sta. Cruz**

IRRI: International Rice Research Institute

**Joie Ramos**

IRRI: International Rice Research Institute

**Gem Faustino**

IRRI: International Rice Research Institute

**Sankalp Bhosale**

IRRI: International Rice Research Institute

**Waseem Hussain** (✉ [waseem.hussain@irri.org](mailto:waseem.hussain@irri.org))

International Rice Research Institute <https://orcid.org/0000-0002-6861-0193>

---

**Original article**

**Keywords:** Rice, Drought breeding, Historical data, Genetic trends, Breeding panel

**Posted Date:** September 22nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-910539/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

1 **Genetic Trends Estimation in IRRIs Rice Drought Breeding Program and Identification of**  
2 **High Yielding Drought-tolerant Lines**

3 Apurva Khanna<sup>1</sup>, Mahender Anumalla<sup>1</sup>, Margaret Catolos<sup>1</sup>, Jérôme Bartholomé<sup>1,2</sup>, Roberto  
4 Fritsche-Neto<sup>1</sup>, John Damien Platten<sup>1</sup>, Daniel Joseph Pisano<sup>1</sup>, Alaine Gullés<sup>1</sup>, Ma Teresa Sta.  
5 Cruz<sup>1</sup>, Joie Ramos<sup>1</sup>, Gem Faustino<sup>1</sup>, Sankalp Bhosale<sup>1</sup>, Waseem Hussain<sup>1\*</sup>

6 <sup>1</sup>Rice Breeding Innovation Platform, International Rice Research Institute (IRRI), Los Baños,  
7 Laguna 4031, Philippines.

8 <sup>2</sup>CIRAD, UMR AGAP, Montpellier, France, AGAP, Univ Montpellier, CIRAD, INRA,  
9 Montpellier SupAgro, Montpellier, France.

10

11 **\* Corresponding author:**

12 Waseem Hussain

13 International Rice Research Institute (IRRI),

14 DAPO Box 7777, Metro Manila, Philippines

15 Email: [waseem.hussain@irri.org](mailto:waseem.hussain@irri.org)

16

## 17 **Abstract**

### 18 **Background**

19 Estimation of genetic trends using historical data is an important parameter to check the success  
20 of the breeding programs. The estimated genetic trends can act as a guideline to target the  
21 appropriate breeding strategies and optimize the breeding program for improved genetic gains. In  
22 this study, 17 years of historical data from IRRI's rice drought breeding program was used to  
23 estimate the genetic trends and assess the success of the breeding program. We also identified  
24 top-performing lines based on grain yield breeding values as an elite panel for implementing  
25 future population improvement-based breeding schemes.

### 26 **Results**

27 A two-stage approach of pedigree-based mixed model analysis was used to analyze the data and  
28 extract the breeding values and estimate the genetic trends for grain yield under non-stress,  
29 drought, and in combined data of non-stress and drought. Lower grain yield values were  
30 observed in all the drought trials. Heritability for grain yield estimates ranged between 0.20-0.94  
31 under the drought trials, and 0.43-0.83 under non-stress trials. Under non-stress conditions the  
32 genetic gain of 0.44% (21.20 kg/ha/year) for genotypes and 0.17 % (7.90 kg/ha/year) for checks  
33 was observed. The genetic trend under the drought conditions exhibited a positive trend with the  
34 genetic gain of 0.11% (1.98kg/ha/year) for genotypes and 0.55% (9.52kg/ha/year) for checks.  
35 For combined analysis showed a genetic gain of 0.39% (12.13 kg/ha/year) for genotypes and  
36 0.60% (13.69 kg/ha/year) for checks was observed. For elite panel selection, 200 promising lines  
37 were selected based on higher breeding values for grain yield and prediction accuracy of >0.40.  
38 The breeding values of the 200 genotypes formulating the core panel ranged between 2366.17  
39 and 4622.59 (kg/ha).

### 40 **Conclusions**

41 A positive genetic rate was observed under all the three conditions; however, the rate of increase  
42 was lower than the required rate of 1.5% genetic gain. We propose a recurrent selection breeding  
43 strategy within the elite population with the integration of modern tools and technologies to  
44 boost the genetic gains in IRRI's drought breeding program. The elite breeding panel identified  
45 in this study forms an easily available and highly enriched genetic resource for future recurrent  
46 selection programs to boost the genetic gains.

47 **Key words: Rice, Drought breeding, Historical data, Genetic trends, Breeding panel**

## 48 **Background**

49 Rice (*Oryza sativa L.*) is one of the world's major staple crops providing up 20% of the world's  
50 dietary energy and feeding more than 3.5 billion people in the world (Wing et al. 2018).  
51 Globally, rice is cultivated in an area of 162Mha with an annual production of 755 mt (FAO  
52 2019). Amongst the area under rice cultivation, more than 30% is under rainfed ecosystems that  
53 are subject to severe drought or water-limited conditions (Dixit et al. 2014). Drought is the major  
54 limitation for rice production in rainfed ecosystems leading to yield loss of 13-35% every year  
55 (Muthu et al. 2020) and affecting 46Mha of rainfed lowland and 10Mha of upland rice  
56 ecosystems in the Asian-Pacific region (Pandey et al. 2007). In Sub-Saharan Africa, drought  
57 covers 19% of the total cultivated rice area and is one of the major causes of low rice grain yields  
58 (Van Oort and Zwart, 2018). Low grain yield under drought conditions is further elevated by the  
59 pressing climatic changes due to the increasing frequency of drought severity events, thus further  
60 limiting the rice productivity (Lenaerts et al. 2019).

61 Direct selection for grain yield under drought has been a major focus of the Rainfed Rice  
62 Breeding (RRB) program at the International Rice Research Institute (IRRI). Direct selection for  
63 grain yield over secondary traits under drought has been proven effective in improving drought  
64 tolerance, and as a result, many drought-tolerant rice varieties have been developed (Kumar et al.  
65 2014; Sandhu and Kumar, 2017; Kumar et al. 2018; Bhandari et al. 2020; Dar et al. 2020).  
66 However, due to the complex nature of grain yield under drought; mainly characterized by the  
67 small and large effect genes; their epistatic interactions, and interaction with environment; and  
68 other abiotic stresses, genetic improvement in drought has been a major challenge.

69 Despite these challenges, IRRI has been constantly striving to innovate and develop drought-  
70 tolerant rice varieties and disseminate them to farmers in Africa and Asia-Pacific regions.  
71 STRASA (Stress Tolerant Rice for Africa and South Asia) Project (2005-2019) at IRRI, funded  
72 by the Bill and Melinda Gates Foundation (BMGF) was one of the most successful research  
73 programs that led to the successful development and release of more than 30 high-yielding  
74 drought-tolerant rice varieties in Asia and Africa (<https://strasa.irri.org/>). Under this project,  
75 imperative efforts were made to incorporate the major drought-tolerant QTLs (*qDTY1.1*,  
76 *qDTY2.1*, *qDTY2.2*, *qDTY3.1*, *qDTY4.1*, *qDTY12.1* *qDTY6.3*, *etc.*,) into the background of the  
77 mega rice varieties like IR64, Swarna, and TDK-1 which led to the development of several high

78 yielding drought-tolerant rice varieties (Bernier et al. 2007; Venuprasad et al. 2009; Vikram et al.  
79 2011; Mishra et al. 2013; Yadaw et al. 2013; Sandhu et al. 2014; Kumar et al. 2014; Henry et al.  
80 2015; Dixit et al. 2017; Sandhu and Kumar 2017; Henry et al. 2019; Sandhu et al. 2019;  
81 Bhandari et al. 2020; Dixit et al. 2020; Kumar et al. 2020; Majumder et al. 2021; Sandhu et al.  
82 2021; Yadav et al. 2021). The most popular drought-tolerant rice varieties include - DRR dhan  
83 42, CR Dhan 801, Sahbhagi dhan in India, Sukha dhan 4, Bahuguni dhan 11 in Nepal, and  
84 Katihan 2, Katihan 3, Sahod Ulan 15, Sahod Ulan 19 in the Philippines, Yaenelo 4 in Myanmar,  
85 MPTSA and ATETE in Malawi, CAR 14 in Cambodia, and BRRI Dhan66, BRRI Dhan 71 in  
86 Bangladesh (<https://strasa.irri.org/variatal-releases/drought>). Despite these endeavors, and the  
87 success of the phenotypic selection coupled with marker-based selection strategies, the progress  
88 in the genetic improvement of the drought breeding program has been limited. For example, the  
89 average estimated rate of genetic gain in rice drought breeding programs in IRRI-India ranges  
90 from 0.68% (under non-stress conditions) to 1.8% (severe drought conditions) (Kumar et al.  
91 2021), which is not sufficient to meet future rice demands. Hence, it is crucial to increase rice  
92 productivity at a greater rate to ensure food security and prevent potential food crises in the  
93 future (Peng et al. 2004; Li et al. 2018). Aiming 1.5% or above genetic gain in rice under drought  
94 is a huge challenge and is largely hampered by complex genetics of drought-elevated by extreme  
95 climatic changes and increase in the frequency of drought events, and availability of limited land  
96 to grow rice.

97 To suffice the increasing food demands, it is important to breed drought-tolerant rice varieties  
98 with expected genetic gains. Rice breeders must be smart to implement the advanced tools and  
99 technologies into the existing breeding pipeline and re-design it for accelerating genetic gains.  
100 The Accelerated Genetic Gain in Rice Alliance (AGGRi) project funded by BMGF is one of the  
101 IRRI's ongoing projects aimed at modernizing the IRRI-NARES (National Agriculture Research  
102 Extension System) rice breeding program and accelerate the genetic gain from the current level  
103 of less than 1% to at least 1.5% or above annually.

104 Genetic gain is an important parameter to check the progress and success of the breeding  
105 programs. The genetic gain estimations can be associated with the ongoing breeding program to  
106 target the appropriate breeding strategies and act as a guide to optimize and modernize the rice  
107 breeding program for accelerated genetic gains. The rate of genetic gain in the IRRI's rice  
108 drought breeding program has never been estimated. On the other note, the historical or current

109 elite breeding lines are an important genetic resource that can be directly used in the population  
110 improvement-based breeding programs to improve the genetic gains. Further, integrating the  
111 modern genomic tools and technologies with the population improvement-based breeding  
112 programs using elite lines as a genetic resource will boost the genetic gains (Xu et al. 2017).  
113 However, it is important to select the appropriate lines from the historical breeding pool  
114 representing the overall genetic diversity in the breeding pool and should have high mean  
115 performance for grain yield and possess the major genes or haplotypes for mendelian traits.

116 Thus, to assess genetic gains in the IRRI's rice drought breeding program and select the valuable  
117 elite lines as a future genetic resource we conducted this study to 1) estimate the genetic trends  
118 for grain yield in IRRI's rice drought breeding program by leveraging 17 years of historical data  
119 from the advanced yield trials (AYT) managed under drought and normal conditions and, 2)  
120 identify high yielding drought-tolerant lines based on the grain yield breeding values as a future  
121 genetic resource for recurrent selection program. The AYT trials included premium released  
122 varieties from IRRI globally, therefore, are the potential reservoirs to select some of the high-  
123 yielding drought-tolerant lines as a future genetic resource.

## 124 **Materials and Methods**

### 125 **Description and Pre-processing of Historical Data**

126 For this study, historical data from yield trials conducted under normal (non-stress) and drought  
127 conditions (at the reproductive stage) at IRRI, Philippines from 2003 to 2019 (17 years) were  
128 used. The trials in each year were conducted in two seasons- dry (from January to April) and wet  
129 (from late June to September) season. The combination of year, season and treatment were  
130 treated as a trial or environment. In total the historical data harbored 19,916 data points with  
131 2,497 unique lines. The data was pre-processed, and the quality of phenotype records was  
132 checked initially to ensure high-quality trials and phenotypes are retained for downstream  
133 analysis. The data was checked for extreme or unexpected values, missing percentages, and valid  
134 experimental design. The trials having more than 20% of missing data for grain yield, lack of  
135 replications, or proper experimental design were dropped initially. Further, the extreme  
136 observations were checked in the data before outlier detection as they may increase the error  
137 variance which may affect the performance of the outlier detection (González et al. 2018). After  
138 preprocessing data was checked for outlier detection using the Bonferroni-Holm test (Bernal-

139 Vasquez et al. 2016; Philipp et al. 2019). The 88 data points detected as outliers were removed  
140 from the data to make sure only high-quality data points are retained for reliable estimates. In  
141 total 53 trials harboring 19,828 data points with 2,490 unique lines were retained for downstream  
142 analysis. The complete information on the trials including the year, season, treatment,  
143 experimental design, number of plots, replications, and blocks are provided in the Additional file  
144 1: Table S1. Trials were performed in varied experimental designs including alpha lattice,  
145 augmented randomized complete block, and randomized complete block designs (RCBD). Three  
146 major agronomic traits days to 50% flowering (DTF), plant height (PH), and grain yield (kg/ha)  
147 were retrieved and used for downstream analysis.

### 148 **Pedigree Data Extraction**

149 The pedigree data of 2,490 unique lines were extracted from the breeding data management  
150 system (Breeding 4 Results (B4R), 2021) which has the passport to comprehensive information  
151 of the genotype, phenotype, and pedigree data of breeding lines. Additional information on the  
152 parents and grandparents up to seven generations, cross-type for each line, and breeding strategy  
153 were extracted using the IRRI genealogy management system (McLaren et al. 2005; Collard et  
154 al. 2019) using customized R scripts.

### 155 **Statistical Modelling of Phenotypic Data**

156 A two-stage approach of mixed model analysis was used to analyze the data and extract the  
157 breeding values for grain yield, DTF, and PH (Piepho et al. 2008; Piepho et al. 2012; Smith and  
158 Cullis 2018) under non-stress, drought, and by analyzing drought and non-stress together. The  
159 two-stage approach was adopted to account for different experimental designs across the  
160 environments (Damesa et al. 2017). In the first stage, per year adjusted means as BLUEs for each  
161 genotype were estimated for each environment. The mixed model used consisted of genotypes as  
162 fixed effects, and season replications and or blocks were used as random effects. BLUEs for each  
163 genotype per year was obtained using the following linear mixed model:

$$y_{ijkl} = \mu + g_i + r_j + b_k + s_l + \varepsilon_{ijkl} \quad (1)$$

164 where,  $y_{ijkl}$  represents adjusted means for  $i$ th observation in  $j$ th replication,  $k$ th block and  $l$ th  
165 season,  $\mu$  is the overall mean,  $g_i$  is the fixed effect of  $i$ th genotype,  $r_j$  is the random effect of  
166 replications in each trial,  $b_k$  is the random block effect,  $s_l$  is the random effect for season and



167  $\varepsilon_{ijkl}$  is the residual error. The random effects were distributed independently and identically. In  
 168 this model, DTF was used as a covariate for reducing the error on yield caused due to the  
 169 presence of different maturity genotypes. The above model was used for the trials which were  
 170 performed using an alpha-lattice breeding design. For environments with augmented RCBD  
 171 experimental design, replications were considered equal to blocks, and hence block effect was  
 172 removed. Likewise, BLUEs and standard error values were calculated respectively for each  
 173 genotype per year for the DTF and PH traits using equation 1.

174 The combined analysis using a linear mixed model was used to extract the single value BLUEs  
 175 adjusted across the non-stress and drought treatments. The model used follows as:

$$y_{ijklm} = g_i + r_j + b_k + s_l + t_m + \varepsilon_{ijklm} \quad (2)$$

176 all the terms are described in equation 1 except the  $t_m$  which is the fixed effect of  $m$ th treatment  
 177 (non-stress and drought). We assume different variances across non-stress and drought  
 178 treatments in the model to get the adjusted means.

179 Heritabilities for grain yield in each environment across non-stress and drought conditions were  
 180 calculated for each environment (Piepho and Mohring 2007). The same model described above  
 181 was used to calculate the heritability with genotypes as a random effect. The equation to  
 182 calculate heritability follows as:

$$H^2 = 1 - \frac{\bar{V}_{BLUP}}{2\sigma_g^2} \quad (3)$$

183 where,  $\bar{V}_{BLUP}$  is the mean-variance difference of two BLUPs and  $\sigma_g^2$  is the variance of  
 184 genotypes.

185 In the second stage analysis, a pedigree-based mixed model approach was used to extract the  
 186 breeding values each in non-stress, drought, and combined data using weighted BLUEs as  
 187 response variable (Mohring and Piepho 2009). The weighted BLUEs were used to take care of  
 188 the heterogeneous error variance and weights were calculated by the inverse of the squared  
 189 standard error of BLUEs. The model used was:

$$y_{ij} = \mu + g_i + ye_j + \varepsilon_{ij} \quad (4)$$

190

191 where  $y_{ij}$  is the weighted BLUE value for  $i$ th observation in  $j$ th year,  $\mu$  is the overall mean,  $g_i$  is  
192 the random effect of  $i$ th genotype with  $g_i \sim N(0, A\sigma_g^2)$  where  $\sigma_g^2$  is the genetic variance and  $A$  is  
193 the additive genetic relationship matrix based on pedigrees,  $ye_j$  is the random effect of year, and  
194  $\varepsilon_{ij}$  is the residual error, with  $\varepsilon_{ij} \sim N(0, R\sigma_\varepsilon^2)$ , where  $R$  is the identity error covariance matrix and  
195  $\sigma_\varepsilon^2$  is the error variance. The reliability of the breeding values (Isik et al. 2017) of each genotype  
196 was estimated using the following equation:

$$r = 1 - \frac{PEV}{\sigma_g^2} \quad (5)$$

197 Two-stage mixed model data analysis was performed in the R software (R Core Team, 2020).  
198 using the ASReml-R 4 package (Butler et al. 2017). The R package AGHMatrix was used for  
199 constructing the pedigree A-matrix (Amadeu et al. 2016). The analytical pipeline and codes are  
200 available on the GitHub ([https://github.com/whussain2/Genetic\\_Trend\\_Rice\\_Drought](https://github.com/whussain2/Genetic_Trend_Rice_Drought))

### 201 **Estimation of the Genetic Trends**

202 The genetic gain was estimated separately for three conditions: a) non-stress (normal conditions)  
203 trials, b) drought trials, and c) combined data (adjusted means across non-stress and stress trials).  
204 For the genetic gain trend, breeding values were regressed on the year of origin of the line. The  
205 genetic trend was also estimated for released varieties and checks by regressing the breeding  
206 value of checks on the year of origin in non-stress trials, drought trials, and combined data.

### 207 **Identification of Breeding Panel**

208 Adjusted breeding values obtained from mixed-model analysis across non-stress and stress data  
209 were used for the identification of elite genotypes as a future breeding resource. A total of 200  
210 lines were selected from the 2,490 unique historical lines based on the higher breeding values  
211 and prediction accuracy of  $>0.4$ . In addition to the lines with higher breeding values and  
212 prediction accuracy for grain yield, lines with key QTLs for key biotic and abiotic stresses. To  
213 make sure genotypes selected are diverse and represent the whole collection of lines in historical  
214 lines, the pedigree matrix was used in the analysis to account for similarity among the lines. The  
215 similarity and diversity among the selected lines in comparison to the whole collection were  
216 visualized through bi-plot. For bi-plot, principal component analysis (PCA) was performed on  
217 the pedigree-based relationship matrix using the *princomp()* function in R software. Bi-plot was  
218 visualized using the *factoextra* R package (Kassambara and Mundt, 2017)

## 219 **Results**

### 220 **Descriptive Features of Historical Drought Data**

221 The three main traits grain yield, PH, and DTF grown under normal and drought conditions were  
222 used for analysis. The difference in phenotypic trait values for all three traits was observed in  
223 normal and drought conditions. The difference in trait value was also evident across the years.  
224 The raw mean grain yield under the normal conditions ranged from 3000-12000 (kg/ha) as  
225 compared to the drought conditions 1360-5600 (kg/ha) (Fig. 1a). In each trial, lower yield values  
226 were observed under drought conditions as compared to trials under normal conditions indicating  
227 the impact of drought on the phenotypes. For DTF we observed higher values under drought  
228 conditions (63-129 days) as compared to the trials under normal conditions (60-119 days)  
229 (Kazan and Lyons 2016; Shavrukov et al. 2017). Further, we observed a wide distribution in  
230 DTF for the genotypes, and genotypes were classified into three maturity groups i.e., early (85-  
231 109 days), medium (110-124 days), and late (DTF>125 days) (Additional file 1: Fig. S1). The  
232 DTF of 68% of the lines from the complete unique set of lines under drought implication falls  
233 into the medium duration maturity category. The remaining genotypes make up 15% and 16%  
234 for early and late duration groups, respectively. Similarly, under the non-stress conditions, the  
235 percentage of lines falling into three maturity categories were, early (15%), medium (76%), and  
236 late (9%) respectively. Because of the wide distribution in DTF, we used DTF as covariate in the  
237 phenotypic data modeling to adjust for the grain yield. Similarly, for PH we observed a wide  
238 distribution in phenotypic values and PH in the data set ranged from 40-195 (cm) under normal  
239 conditions and 40-90 (cm) under drought conditions (Additional file 1: Fig. S2). The dataset has  
240 ample diversity among the tested genotypes having a diverse range for PH between 40 cm to 195  
241 cm. Low plant height was observed under drought conditions, consistent with the previous  
242 reports (Ahmadikhah and Marufinia 2016; Mishra and Panda 2017; Hussain et al. 2018; Panda et  
243 al. 2021). Heritability for grain yield estimates ranged between 0.20-0.94 under the drought trials  
244 and 0.43-0.83 under normal growing trials for the non-stress trials (Fig. 1b). Lower heritability  
245 was observed in most of the drought trials. Reduction in heritability under drought conditions is a  
246 common phenomenon (Henry et al. 1997, Kumar et al. 2007) which indicates that genotypes are  
247 not able to express the higher genetic potential for grain yield.

### 248 **Historical Data Connectivity**

249 Historical data sets usually have very low connectivity as new lines are being tested every year,  
250 and only a limited number of times the new lines are being tested. In the current data set, we  
251 observe appropriate connectivity of the different genotypes across the years (Fig. 1c), and this  
252 connectivity was mainly created by long-term checks (IR64, Swarna, Sahbhagi Dhan, IRRI 154)  
253 across the years. Further, over the breeding cycles and years, the superior genotypes were  
254 forwarded and re-tested in the succeeding years which made the dataset well connected to  
255 previous years (Additional file 1: Fig. S3) Further to ensure good connectivity and get reliable  
256 estimates of the breeding values, relationship matrix (Additional file 1: Fig. S4) based on  
257 pedigrees of 2,970 unique lines was incorporated in the second stage of mixed model analysis.

### 258 **Estimation of Breeding Values**

259 Breeding values obtained from second-stage analysis by fitting a pedigree matrix were used to  
260 estimate the genetic trends and used to identify the best lines based on higher breeding values for  
261 the formulation of the core panel. The range of the breeding values for the genotypes under  
262 drought stress was between 642.79- 3,267.60 (kg/ha). Under the normal growing conditions, the  
263 breeding values of the genotypes ranged between 3,447.93- 6,933.32 (kg/ha). The breeding  
264 values adjusted across drought and normal growing conditions ranged between 1,026.93-  
265 4,622.59 (kg/ha). The histogram of the breeding values along with the mean and standard  
266 deviations is given in the (Additional file 1: Fig. S5).

### 267 **Estimation of Genetic Trends**

268 The genetic trend was estimated for the genotypes and, also for checks and released varieties in  
269 normal growing conditions, drought conditions, and in combined data. Under normal growing  
270 conditions, the genetic gain of 0.44% with a yield advantage of 21.20 kg/ha per year was  
271 observed for genotypes (Fig. 2a), and genetic gain of 0.17% was observed for checks and  
272 released lines representing increase of 7.90 kg/ha per year (Fig. 3a). The genetic trend under the  
273 drought conditions exhibited a positive trend with a genetic gain of 0.11% for genotypes (Fig.  
274 2b) and 0.55% for released lines (Fig. 3b). Yield advantages of 1.98 kg/ha for genotypes and  
275 9.52 kg/ha for checks was observed. The regression estimates for combined analysis (adjusted  
276 breeding values across normal and drought growing conditions) showed a genetic gain of 0.39%  
277 for genotypes (Fig. 2c) and 0.60% (Fig. 3c) for checks with a yield advantage of 12.13 kg/ha for  
278 genotypes and 13.69 kg/ha for checks.

279 Further we divide the genetic gain trend estimation in time periods of 2003-2012 and 2013-2019  
280 based on the breeding strategies adopted by the drought breeding program at IRRI. Genetic gain  
281 was 0.31 % (under non-stress condition) and 0.21 % (under stress condition) higher in the period  
282 2013-2019 as compared to period 2003-2012 (Additional file 1: Fig. S6)

### 283 **Comparison of Breeding Values**

284 We compared the breeding values of the popular checks and the released varieties from IRRI, to  
285 assess their performance under three conditions *viz.*, stress, non-stress, and combined stress and  
286 non-stress. This assessment was undertaken with the overview to deduce the best performers  
287 among the released IRRI lines. The performance of the released varieties was superior to the  
288 popular checks under all three conditions (Fig. 3a-c). The varieties IRRI 188, IRRI 199, and  
289 IRRI 200 had superior breeding values of 2,547.19 (kg/ha), 2,420.70 (kg/ha), and 2,490.97  
290 (kg/ha), respectively, compared to popular checks and other released varieties when estimated  
291 under the drought stress conditions (Fig. 3a). These varieties were released in the years 2015 and  
292 2016, depicting growth of the breeding program over the preceding released varieties developed  
293 in the previous years. The latter also showed the higher performance to the most popular  
294 drought-tolerant released variety Sahbhagi Dhan with the predicted value of 1,562.03 (kg/ha)  
295 under stress conditions. The breeding values for popular drought-tolerant checks Vandana and  
296 Rajashree were 1,391.68 (kg/ha) and 1815.65 (kg/ha) under drought stress conditions  
297 respectively. This depicts the superiority of the recently released lines over popular checks and  
298 varieties. The genotypes were also assessed for the trials evaluated under non-stress conditions.  
299 The breeding lines, IR09L204 depicted the highest breeding value of 4998.44 (kg/ha) followed  
300 by IRRI 199 (4,958.74), IRRI 218 (4,900.64), IRRI 200 (4,892.13), and IRRI 188 (4,866.45) in  
301 the descending fashion of the superior performing genotypes under the non-stress conditions  
302 (Fig. 3b). In the combined analysis with breeding lines performing superior under both drought  
303 stress and non-stress conditions were as follows IRRI 119 (3,698.75), IRRI 163 (3,599.56), IRRI  
304 162 (3,395.07), IRRI 218 (3,374.91), and IRRI 200 (3374.93) (Fig. 3c). Similar to stress  
305 conditions, these lines surpassed popular checks Vandana (2,814.09) and Rajashree (2,628.67);  
306 variety Sahbhagi Dhan (2,668.12). The breeding lines of the five top-performing lines have been  
307 detailed in Fig. 3a-c. for stress, non-stress, and combined conditions, respectively.

### 308 **Identification and Development of Elite Breeding Panel**

309 The genotypes with higher breeding values based on the grain yield and with prediction accuracy  
310 of >0.40 were selected for the development of the breeding core panel. In total 200 promising  
311 lines were identified and used for the development of the elite panel. To make sure 200 selected  
312 lines are genetically diverse and representative of decades of IRRI's drought breeding and  
313 varietal development, a pedigree matrix was used to account for genetic similarity among the  
314 genotypes. The relationship matrix constructed using pedigree data was visualized using the  
315 biplot (Fig. 3d). From the biplot, it is very clear how diverse the selected lines are, and how they  
316 represent and capture the diversity of the whole historical collection of 2,490 unique genotypes.  
317 The breeding values of the 200 genotypes formulating the core panel ranged between 3,200 to  
318 4,622.59 kg/ha. The mean breeding value of the panel is 3,395.10 with a standard deviation of  
319 3,73.23 (Additional file 1: Fig. S7).

## 320 **Discussion**

321 Here we provide an overview of how 17 years of historical IRRI's rice drought data was  
322 leveraged to estimate the breeding values for grain yield and estimate the genetic trends in the  
323 IRRI's rice drought breeding program. We also demonstrated how the top-performing lines  
324 based on the grain yield breeding values were selected as the future breeding elite resource for  
325 recurrent selection-based breeding. For the selection of genotypes as a part of the core panel,  
326 pedigree information availability was pivotal in fitting the additive matrix in the second stage of  
327 the mixed model analysis for reliable estimation of the breeding values and help in the selection  
328 of accurate genotypes for the formulation of the core panel. The essence of using the relationship  
329 matrix is that it contains information about the flow of genes and explicitly allows the dissection  
330 of genetic variation by accounting for the additive genetic covariances between random  
331 effects/genotypes for reliable estimation of the breeding values (Piepho et al. 2008). Further, the  
332 relationship matrix ties up the data across years by borrowing information from parents and  
333 grandparents and creating the connectivity in the highly unbalanced data set for the reliable  
334 estimation of breeding values.

## 335 **Genetic Trend and Breeding Value Estimations**

336 Improving the crop yield or genetic gain is crucial for minimizing the "yield gap" between the  
337 breeders and farmers. The assessment of genetic gains to estimate crop yield growth has had a  
338 limited focus in the past. However, genetic gain estimation with an outlook to reinforce the

339 future breeding programs to increase genetic gain for yield has become a major focus. In this  
340 study, the regression of breeding values over the year of release/testing indicated a positive  
341 genetic gain under all three conditions, and the overall success of the drought breeding program  
342 at IRRI. However, the genetic gain of 0.11% observed under drought conditions is not sufficient  
343 to meet the current and future rice food demands. A much higher genetic gain of 1.9% was  
344 reported under severe drought conditions at the reproductive stage in rice evaluated in IRRI India  
345 (Kumar et al. 2021). Therefore, it is essential to optimize and modernize the drought rice  
346 breeding program at IRRI for enhanced genetic gains.

347 Further we observed a minimal increase in genetic gain increase till the year 2012, post which  
348 there was higher increase in genetic gain (Additional file 1: Fig.S6) which could be accredited to  
349 differential breeding strategies followed in the drought breeding program across the years. The  
350 drought breeding program at IRRI until 2012 was led specifically for targeting the introgression  
351 of major abiotic stress-tolerant QTLs/genes for the development of NILs possessing elite genetic  
352 backgrounds with minimal focus on recurrent selection breeding strategy. Thereafter, in the  
353 preceding years, the focus drifted towards pyramiding these genes/QTLs for rendering multiple  
354 stress-tolerant cultivars. These genotypes were not precisely targeted for genetic gain  
355 enhancement; however, their genetic merit is highly valuable as these sustainable varieties were  
356 disseminated for commercial cultivation across countries and can withstand additional biotic and  
357 abiotic stresses along with drought. Further in the last few years more emphasize was given to  
358 recurrent selection-based breeding strategies in the drought breeding program to improve the  
359 yield and increase genetic gain. Rather than the QTL-based introgression or breeding approach  
360 as used previously, rice breeders need to focus on population improvement breeding approaches  
361 using elite lines as parents, wherein parents of each breeding cycle are selected based on high  
362 additive breeding values for the grain yield. Recurrent selection schemes focused on quick  
363 recycling of the best and high breeding value lines may deliver higher rates of genetic gain  
364 (Cobb et al. 2019). Further to ensure the constant genetic gains, rice breeders need to select the  
365 parents for new breeding cycles that have higher additive breeding value than the previous  
366 breeding cycle. However, for long-term genetic gains, utmost care must be taken by the breeders  
367 to diversify the elite gene pool by bringing or directly crossing the exotic or diverse materials  
368 with elite pool lines. Exotic or diverse materials may broaden the elite gene pool, but they are  
369 highly unadopted and unimproved lines with lower breeding values. Thus, crossing a directly

370 diverse line with an elite line may bring novel favorable alleles and increase the genetic variance  
371 of progeny, however, they barely counterbalance the mean performance of progeny due to the  
372 low breeding value of diverse or exotic lines (Longin et al. 2014; Allier et al. 2020). Thus, a  
373 focused recurrent selection breeding approach with a systematic pre-breeding approach is  
374 required to deliver higher and constant genetic gains in IRRI's drought rice breeding program.

375 The drought breeding program at IRRI has successfully released many drought-tolerant varieties  
376 across Asia and Africa (<https://strasa.irri.org/>), and most of them were part of this study. We  
377 separately assessed the genetic gain estimation for released varieties, and higher genetic gain was  
378 observed for varieties/checks released for use under drought conditions. The positive and higher  
379 genetic gain (0.55%) for checks under drought indicates that a strong impact has been made by  
380 IRRI's drought breeding program to increase rice productivity under challenging and extreme  
381 environments. Among these varieties- IRRI 188, IRRI 199, and IRRI 200 released during 2015  
382 and 2016 had higher breeding values as compared to the popular checks Rajashree, Vandana, and  
383 UPLRi7, and previously released varieties. The superiority of these lines over the formerly  
384 released varieties indicates growth in grain yield to a larger extent. Also, positive genetic trend  
385 and higher breeding values of the recently released varieties demarcates the positive growth of  
386 the breeding program across years as the performance improved from the preceding released  
387 varieties developed in the previous years. Further, among these three top-performing released  
388 lines, IRRI 199 originates from a backcross breeding program utilizing a tropical japonica  
389 drought and blast tolerant genotype, Moroberekan, as a donor parent and a high-yielding, semi-  
390 dwarf Indica rice variety, Swarna. The population harbors a major, severe drought-tolerant QTL,  
391 *qDTY3.2* contributing to various drought-tolerant traits *viz.*, canopy cover, canopy temperature  
392 (CT), root system architecture (RSA) attributes (Wasaya et al. 2018; Sofi et al. 2019). The  
393 genetic region co-localizes with early flowering time QTL *HD9* and lodging resistance features,  
394 making the genotype suitable for various ecosystems and environments (Dixit et al. 2014; 2015).  
395 Concomitantly, based on the combined analysis IRRI 163, IRRI 162, IRRI 218 and IRRI 200  
396 released in the years 2011, 2015, and 2016 exhibited high breeding values under both stress and  
397 non-stress conditions.

### 398 **Development of Breeding Panel as The Future Breeding Resource**



399 In the last one-decade rice breeders at IRRI have mainly focused on introgression and  
400 pyramiding of major abiotic stress-tolerant QTLs/genes in elite backgrounds (Venuprasad et al.  
401 2009; Mishra et al. 2013; Yadav et al. 2013). Population improvement based on recurrent  
402 selection and early re-cycling of advanced lines has not been a major focus of the drought rice  
403 breeding program at IRRI. Different crossing strategies (Additional file 1: Fig. S8) single,  
404 complex, double, and backcrosses have been used by the rice breeders to integrate these QTLs  
405 into the elite genetic backgrounds and develop the new breeding lines. Diverse materials,  
406 including landraces, and donors have been extensively used to diversify the gene pool and  
407 develop climate-resilient varieties (Sandhu et al. 2021, Yadav et al. 2021). However, this strategy  
408 of diversifying the elite gene pool with limited focus on recurrent selection and early re-cycling  
409 of high-value breeding lines may have limited genetic gain to a large extent and has not been  
410 sufficient to maintain the higher genetic gains over time.

411 Recurrent selection with early re-cycling of lines is the key to increase the frequency of desirable  
412 additive haplotypes of grain yield in each cycle, and ultimately boosting the genetic gains. To  
413 strictly focus on recurrent selection breeding schemes, the presence of highly characterized elite  
414 lines with higher breeding values for grain yield and possessing the key haplotypes for  
415 mendelian traits is required as the base population. The historical data set used in this study  
416 which contains 2,497 unique genotypes has been used by rice breeders at IRRI for decades and  
417 in the past 60 years, many promising drought lines have been extracted from this breeding pool.  
418 This breeding pool exhibits ample genetic diversity and possesses the key lines that may be used  
419 as a future breeding resource to sustain higher grain yield under challenging environments.  
420 Further, this breeding collection has not only been improved for grain yield but also  
421 turbocharged with diverse alleles for important traits of biotic and abiotic stresses and represents  
422 the overall diversity and breadth of IRRI's rice drought breeding program. Besides the gene bank  
423 resources, this breeding pool represents the important source of genetic variation that is highly  
424 dynamic created through recombination and reshuffling of alleles. Thus, identifying the high-  
425 power performing lines based on the breeding values for grain yield that represents the overall  
426 diversity of the whole breeding collections is the key to success in future recurrent selection  
427 breeding strategies. To this end, we took this initiative to extract the top breeding lines from the  
428 whole historical breeding collection and form the elite breeding panel as future breeding  
429 resources. These selected lines besides possessing high breeding values are also indicative of

430 higher recovery capabilities under drought stress. We believe the lines selected are the best  
431 genetic variation to recombine and reshuffle in recurrent selection to increase the frequency of  
432 additive haplotypes of grain yield in each cycle of breeding. Additionally, based on the literature  
433 we assessed the additional characteristic features of these selected lines for resistance to biotic  
434 stresses (bacterial blight, blast, brown planthoppers, stemborer, whiteheads, green leafhoppers,  
435 and rice tungro virus), drought tolerance, and quality traits (Additional file 1: Table S2). It is  
436 evident from the table that these lines harbor the key genes and is a readily available elite genetic  
437 resource for future recurrent selection breeding schemes for targeting preferred  
438 environments/countries based on the desirable market profiles.

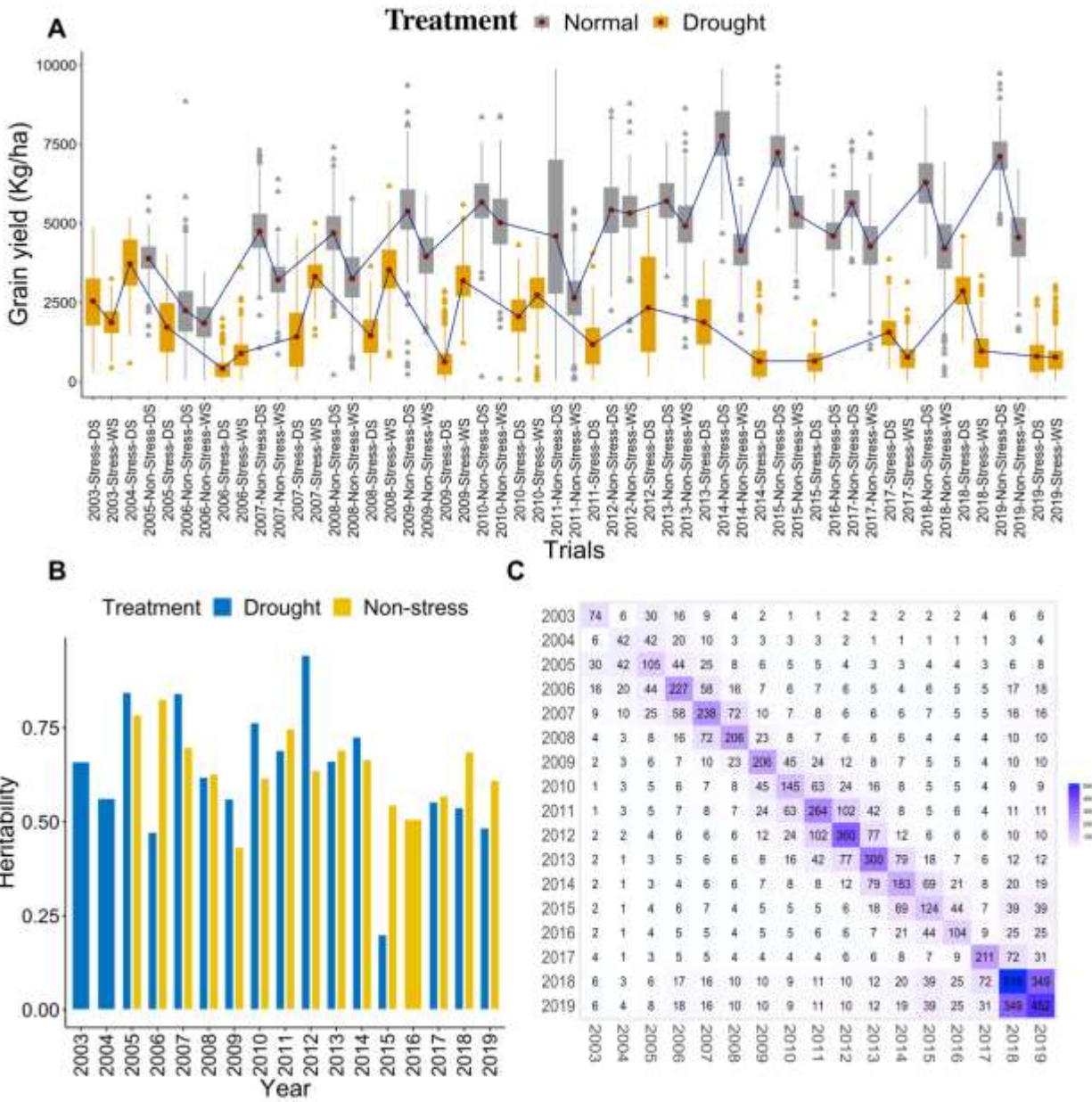
439 Additionally, the elite panel also harbored two best performing genotypes; IR15F1706 and IR  
440 54447-3B-10-2 from the 2020 drought stress trials (unpublished) showing high chlorophyll  
441 fluorescence (CF) and low CT values. These genotypes when assessed had high breeding values,  
442 confirming further the accuracy of the analysis undertaken. Alongside, it also demarcates that  
443 deprivation in CT has a strong influence on plant's yield under drought-prone conditions. It has  
444 also been reported that CT has a high correlation with the RSA traits symbolizing enhanced  
445 genetic capacity of the plant to retain soil moisture and hence improved survival and yield under  
446 drought stress conditions (Blair et al. 2010; Lopes and Reynolds 2010). Similarly, enhanced CF  
447 demarcates a plant's capabilities to withstand drought stress effectively. Furthermore, two  
448 multiparent conventional bred lines namely, IR 115844-B-B-281-1-2 and IR 115844-B-342-1-1-  
449 1 present in the selected panel have been reported to yield more than 7000 kg/ha under dry  
450 direct-seeded conditions, with higher yields under non-stress and reproductive stage drought  
451 stress conditions (Sandhu et al. 2021). In our study, these lines showed the breeding values  
452 between 3200 to 3368.14, which form a valuable resource for the breeding programs to be  
453 utilized for drought-prone areas with major cultivation under dry seeded conditions having  
454 limited water and labor availability. Few of the top 100 selected lines were also reported by other  
455 research studies to show higher yield performance under multi-environment trials (Vergara et al.  
456 2016).

457 In summary, the core breeding panel selected based on the breeding values and prediction  
458 accuracy is an important genetic resource possessing multiple stress tolerance, varied range of  
459 quality traits with genotypes suited for cultivation under challenging environments. Furthermore,  
460 they form an easily available and highly enriched genetic resource for future recurrent selection

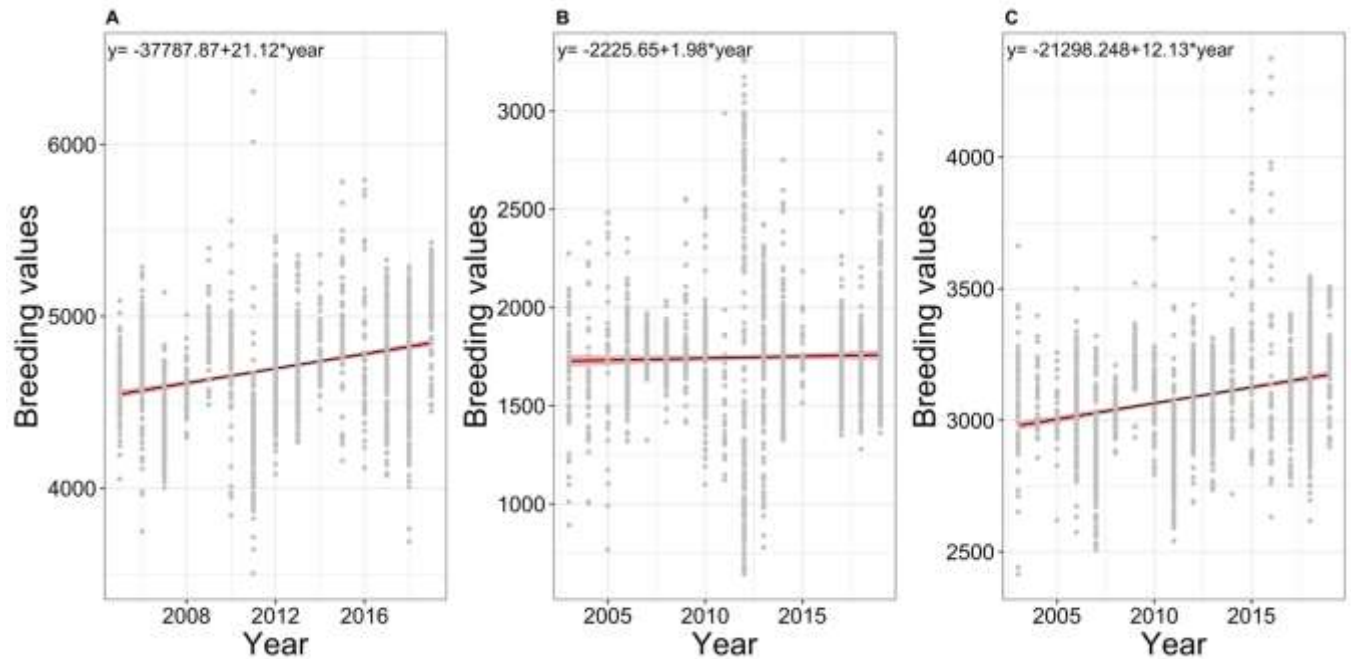
461 programs and enhance genetic gains. However, we emphasize systematic genotypic and  
462 phenotypic characterization of these lines in achieving more knowledge on the value proposition  
463 of these lines, new allele enrichment, and help to create a framework for better understanding  
464 and managing the genetic diversity in the elite pool. However, the question may arise whether  
465 continued use of the elite pool lines and reshuffling of alleles in closed recurrent selection  
466 strategies is enough to maintain long-term genetic gains? Most will agree with the enrichment of  
467 the elite breeding pool with diverse materials as was done previously in the IRRI's drought  
468 breeding program. However, we emphasize here a systematic effort to diversify the elite gene  
469 pool without contaminating it with diverse materials and limiting the genetic gains (Allier et al.  
470 2020).

#### 471 **Conclusion**

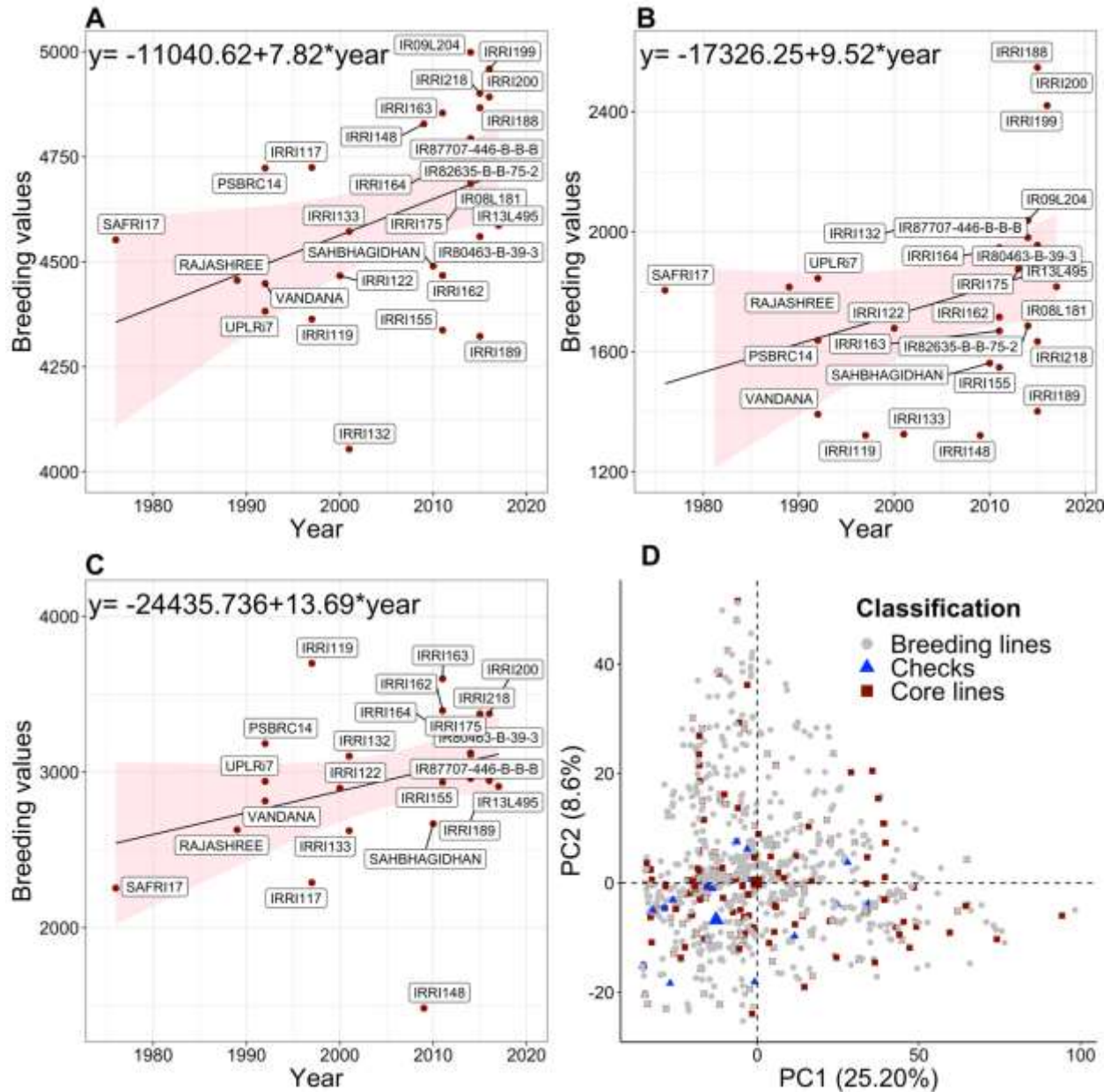
472 The drought breeding program at IRRI has been successful in maintaining a positive genetic rate  
473 in the breeding program, however, the increase in genetic gain has not been so high to fulfill the  
474 rice food demands. To achieve the required genetic gains of 1.5% or above, a recurrent selection  
475 breeding strategy of the elite population with the integration of modern tools and technologies is  
476 needed. Genotypic and phenotypic characterization of the selected elite panel is required to  
477 effectively manage, incorporate, and track the genetic diversity for short-term and long-term  
478 genetic gains. Further, efficient pre-breeding strategies are needed to turbocharge the elite gene  
479 pool with major haplotypes of traits showing discrete Mendelian segregation without  
480 compromising the performance of elite lines and boost the genetic gains.



482  
 483 **Figure 1:** (a) Boxplot showing the mean grain yield (kg/ha) under normal and drought conditions  
 484 from the year 2003 -2019. The x-axis shows the trial names, which are combinations of year,  
 485 season, and growing condition. In the boxplots, it is clear the grain yield is higher under normal  
 486 conditions as compared to drought conditions indicating the impact of drought on the yield trials.  
 487 (b). Heritabilities of the trials in each year from 2003 to 2019. The blue bars represent drought and  
 488 yellow non-stress trials. (c). Connectivity of all the genotypes across years from 2003 to 2019. The  
 489 genotypes including common checks and promising varieties were repeatedly tested for their  
 490 performance in the successive years, thereby making the dataset well connected across successive  
 491 years. The numbers in the boxes show the genotypes that were common between years.



492  
 493 **Figure 2:** Trends in genetic gain from IRRI's 17 years of drought breeding program under. a) Stress  
 494 conditions, b) non-stress conditions, and c) combined conditions (adjusted breeding values under stress  
 495 and non-stress conditions). The x-axis shows the year of origin of the genotype and the Y-axis shows the  
 496 breeding value of the genotype. The genetic gain was estimated by regressing the breeding values of  
 497 grain yield on the year of origin and is given by the slope of the line.  
 498



499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510

**Figure 3:** Trends in genetic gain for IRRIs drought released lines and popular checks under three conditions: a) non-stress, b) reproductive drought and c) combined conditions. The genetic gain was estimated by regressing the breeding values on the year of origin, d) shows the biplot of the lines selected based on the breeding value for grain yield as an elite core panel. Core panel lines are highlighted in dark red color. The checks and released lines are shown in blue color, and the whole historical collection lines are represented in gray color. The biplot was constructed using the first two principal components obtained from the pedigree-based relationship matrix. The selected lines represent and capture the variability of the whole collection of genotypes and are ideal to form the core panel as a future breeding resource.

511 **Availability of Data and Materials**

512 The datasets included in this work are given as additional files. The additional data and scripts  
513 used to run the analysis can be found on the GitHub page at the following link:  
514 <https://github.com/whussain2/Genetic Trend Rice Drought>

515 **Abbreviations**

516 BLUP: Best linear unbiased predictions  
517 BLUE: Best linear unbiased estimates  
518 REML: Residual maximum likelihood  
519 GEBV: Genomic estimated breeding values  
520 CV: Coefficient of variation  
521 AYT: Advanced yield trials  
522 PH: Plant height  
523 DTF: Days to 50% flowering  
524 PCA: Principal component analysis  
525 QTL: Quantitative trait loci  
526 MABB: Marker-assisted backcross breeding  
527 RSA: Root system architecture  
528

529 Ahmadikhah A, Marufinia A (2016) Effect of reduced plant height on drought tolerance in  
530 rice. *3 Biotech* 6:221. <https://doi.org/10.1007/s13205-016-0542-3>

531 Allier A, Teyssèdre S, Lehermeier C, Lehermeier C, Moreau L, Charcosset A (2020)  
532 Optimized breeding strategies to harness genetic resources with different performance  
533 levels. *BMC Genomics* 21:349. <https://doi.org/10.1186/s12864-020-6756-0>

534 Amadeu RR, Cellon C, Olmstead JW, Gracia AA, Resende MF, Muñoz PR (2016)  
535 AGHmatrix: R package to construct relationship matrices for autotetraploid and  
536 diploid species: A blueberry example. *Plant Genome* 9:3.  
537 <https://doi.org/10.3835/plantgenome2016.01.0009>

538 Bernal-Vasquez AM, Utz H-F, Piepho HP (2016) Outlier detection methods for generalized  
539 lattices: a case study on the transition from ANOVA to REML. *Theor Appl Genet*  
540 129:787–804. <https://doi.org/10.1007/s00122-016-2666-6>

541 Bernier J, Kumar A, Ramaiah V, Spaner D, Atlin G (2007) A large-effect QTL for grain yield  
542 under reproductive-stage drought stress in upland rice. *Crop Sci* 47:507–516.  
543 <https://doi.org/10.2135/cropsci2006.07.0495>

544 Bhandari A, Sandhu N, Bartholome J, Hamadoun TV, Ahmadi N, Kumari N, Kumar A (2020)  
545 Genome-wide association study for yield and yield related traits under reproductive  
546 stage drought in a diverse indica-aus rice panel. *Rice* 13:53.  
547 <https://doi.org/10.1186/s12284-020-00406-3>

548 Blair MW, González LF, Kimani PM, Butare L (2010) Genetic diversity, inter-gene pool  
549 introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from  
550 Central Africa. *Theor Appl Genet* 121:237–248. <https://doi.org/10.1007/s00122-010-1305-x>

552 Butler DG, Cullis BR, Gilmour AR, Gogel BG, Thompson R (2017) ASReml-R reference  
553 manual version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

554 Cobb JN, Juma RU, Biswas PS, Juan DA, Rutkoski J, Gary A, Hagen T, Quinn M, Eng NH  
555 (2019) Enhancing the rate of genetic gain in public-sector plant breeding programs:



556 lessons from the breeder's equation. *Theor Appl Genet* 132:627–645.  
557 <https://doi.org/10.1007/s00122-019-03317-0>

558 Collard BCY, Gregorio GB, Thomson MJ, Islam MR, Vergara GV, Laborte AG, Nissila E,  
559 Kretschmar T, Cobb JN (2019) Transforming rice breeding: re-designing the irrigated  
560 breeding pipeline at the international rice research institute (IRRI). *Crop Breed Genet*  
561 *Genom* 1:e190008. <https://doi.org/10.20900/cbgg20190008>

562 Damesa TM, Möhring J, Worku M, Piepho HP (2017) One step at a time: stage-wise analysis  
563 of a series of experiments. *Agronomy* 109:845–857.  
564 <https://doi.org/10.2134/agronj2016.07.0395>

565 Dar MH, Waza SA, Shukla S, Zaidi NW, Nayak S, Hossain M, Kumar A, Ismail AM, Singh  
566 US (2020) Drought tolerant rice for ensuring food security in eastern india.  
567 *Sustainability* 12:2214. <https://doi.org/10.3390/su12062214>

568 Dixit S, Grondin A, Lee C-R, Henry A, Olds TM, Kumar A (2015) Understanding rice  
569 adaptation to varying agro-ecosystems: trait interactions and quantitative trait loci.  
570 *BMC Genet* 16:86. <https://doi.org/10.1186/s12863-015-0249-1>

571 Dixit S, Singh A, Sta Cruz MT, Maturan PT, Amante M, Kumar A (2014) Multiple major QTL  
572 lead to stable yield performance of rice cultivars across varying drought intensities.  
573 *BMC Genet* 15:16. <https://doi.org/10.1186/1471-2156-15-16>

574 Dixit S, Singh UM, Singh AK, Alam S, Challa V, Nachimuthu VV, Yadav S, Abbai R,  
575 Selvaraj R, Devi MN, Ramayya PJ, Badri J, Ram T, Lakshmi J, Lakshmidevi G,  
576 Vidhya LRK, Padmakumari AP, Laha GS, Prasad MS, Seetalam M, Singh VK, Kumar  
577 A (2020) Marker assisted forward breeding to combine multiple biotic-abiotic stress  
578 resistance/tolerance in rice. *Rice* 13:29. <https://doi.org/10.1186/s12284-020-00391-7>

579 Dixit S, Yadaw RB, Mishra KK, Kumar A (2017) Marker-assisted breeding to develop the  
580 drought-tolerant version of Sabitri, a popular variety from Nepal. *Euphytica* 213:184.  
581 <https://doi.org/10.1007/s10681-017-1976-3>

582 FAO (2019) The State of Food and Agriculture 2019. Moving forward on food loss and waste  
583 reduction. Rome. Licence: CC BY-NC-SA 3.0 IGO.

584 Gonzalez MY, Weise S, Zhao Y, Philipp N, Arend D, Börner A, Oppermann M, Graner A,  
585 Reif JC, Schulthess AW (2018) Unbalanced historical phenotypic data from seed  
586 regeneration of a barley ex situ collection. *Sci Data* 5:180278.  
587 <https://doi.org/10.1038/sdata.2018.278>

588 Henry A, Stuart-Williams H, Dixit S, Kumar A, Farquhar G (2019) Stomatal conductance  
589 responses to evaporative demand conferred by rice drought-yield quantitative trait  
590 locus qDTY12.1. *Functional Plant Biol* 46:660–669. <https://doi.org/10.1071/FP18126>

591 Henry A, Swamy BPM, Dixit S, Torres RD, Batoto TC, Manalili M, Anantha MS, Mandal NP,  
592 Kumar A (2015) Physiological mechanisms contributing to the QTL-combination  
593 effects on improved performance of IR64 rice NILs under drought. *J Exp Bot*  
594 66:1787–1799. <https://doi.org/10.1093/jxb/eru506>

595 Henry T, Nguyen R, Chandra Babu, Blum A (1997) Breeding for Drought Resistance in Rice:  
596 Physiology and Molecular Genetics Considerations. *Crop Sci* 37: 1426-1434.  
597 <https://doi.org/10.2135/cropsci1997.0011183X003700050002x>

598 Hussain HA, Hussain S, Khaliq A, Ashraf U, Anjum SA, Men S, Wang L (2018) Chilling and  
599 drought stresses in crop plants: implications, cross talk, and potential management  
600 opportunities. *Front Plant Sci* 9:393 <https://doi.org/10.3389/fpls.2018.00393>

601 Isik F, Holland J, Maltecca C (2017) Spatial Analysis. In: *Genetic Data Analysis for Plant and*  
602 *Animal Breeding*. Springer International Publishing, Cham, pp 203–226

603 Kassambara A, Mundt F (2017) Package “factoextra” for R: Extract and Visualize the Results  
604 of Multivariate Data Analyses. R Package. version 2017, 1–77.

605 Kazan K, Lyons R (2016) The link between flowering time and stress tolerance. *J Exp Bot*  
606 67:47–60. <https://doi.org/10.1093/jxb/erv441>

607 Kumar A, Dixit S, Ram T, Yadav RB, Mishra KK, Mandal NP (2014) Breeding high-yielding  
608 drought-tolerant rice: genetic variations and conventional and molecular approaches. *J*  
609 *Exp Bot* 65:6265–6278. <https://doi.org/10.1093/jxb/eru363>

610 Kumar A, Raman A, Yadav S, Verulkar SB, Mandal NP, Singh ON, Swain P, Ram T, Badri J,  
611 Dwivedi JL, Das SP, Singh SK, Singh SP, Kumar S, Jain A, Chandrababu R, Robin S,

612 Shashidhar HE, Hittalmani S, Satyanarayana P, Venkateshwarlu C, Ramayya J, Naik  
613 S, Nayak S, Dar MH, Hossain SM, Henry A, Piepho HP (2021) Genetic gain for rice  
614 yield in rainfed environments in India. *Field Crops Res* 260:107977.  
615 <https://doi.org/10.1016/j.fcr.2020.107977>

616 Kumar A, Sandhu N, Dixit S, Yadav S, Swamy BM, Shamsudin NAA (2018) Marker-assisted  
617 selection strategy to pyramid two or more QTLs for quantitative trait-grain yield under  
618 drought. *Rice* 11:35. <https://doi.org/10.1186/s12284-018-0227-0>

619 Kumar A, Sandhu N, Venkateshwarlu C, Priyadarshi R, Yadav S, Majumder RR, Singh VK  
620 (2020) Development of introgression lines in high yielding, semi-dwarf genetic  
621 backgrounds to enable improvement of modern rice varieties for tolerance to multiple  
622 abiotic stresses free from undesirable linkage drag. *Sci Rep* 10:13073.  
623 <https://doi.org/10.1038/s41598-020-70132-9>

624 Kumar R, Venuprasad R, Atlin GN (2007) Genetic analysis of rainfed lowland rice drought  
625 tolerance under naturally-occurring stress in eastern India: Heritability and QTL  
626 effects. *Field Crops Res* 103:42-52. <https://doi.org/10.1016/j.fcr.2007.04.013>.

627 Lenaerts B, Collard BCY, Demont M (2019) Review: Improving global food security through  
628 accelerated plant breeding. *Plant Sci* 287:110207.  
629 <https://doi.org/10.1016/j.plantsci.2019.110207>

630 Li H, Rasheed A, Hickey LT, He Z (2018) Fast-forwarding genetic gain. *Trends Plant Sci*  
631 23:184–186. <https://doi.org/10.1016/j.tplants.2018.01.007>

632 Longin CFH, Reif JC (2014) Redesigning the exploitation of wheat genetic resources. *Trends*  
633 *Plant Sci* 19:631–636. <https://doi.org/10.1016/j.tplants.2014.06.012>

634 Lopes MS, Reynolds MP (2010) Partitioning of assimilates to deeper roots is associated with  
635 cooler canopies and increased yield under drought in wheat. *Functional Plant Biol*  
636 37:147–156. <https://doi.org/10.1071/FP09121>

637 Majumder RR, Sakhale S, Yadav S, Sandhu N, Hassan L, Hossain MA, Kumar A. (2021)  
638 Molecular breeding for improving drought tolerance in rice: recent progress and future  
639 perspectives. In: Hossain MA, Hassan L, Ifterkharuddaula KM, Kumar A, Henry R

640 (eds) Molecular breeding for rice abiotic stress tolerance and nutritional quality. John  
641 Wiley & Sons Ltd. Pp 53-74. <https://doi.org/10.1002/9781119633174.ch3>

642 McLaren CG, Bruskiwich RM, Portugal AM, Cosico AB (2005) The international rice  
643 information system. a platform for meta-analysis of rice crop data. *Plant Physiol*  
644 139:637–642. <https://doi.org/10.1104/pp.105.063438>

645 Mishra KK, Vikram P, Yadaw RB, Swamy BM, Dixit S, Sta Cruz MT, Paul M, Marker S,  
646 Kumar A (2013) qDTY12.1: a locus with a consistent effect on grain yield under  
647 drought in rice. *BMC Genet* 14:12. <https://doi.org/10.1186/1471-2156-14-12>

648 Mishra SS, Panda D (2017) Leaf traits and antioxidant defense for drought tolerance during  
649 early growth stage in some popular traditional rice landraces from Koraput, India. *Rice*  
650 *Sci* 24:207–217. <https://doi.org/10.1016/j.rsci.2017.04.001>

651 Möhring J, Piepho HP (2009) Comparison of weighting in two-stage analysis of plant breeding  
652 trials. *Crop Sci* 49:1977–1988. <https://doi.org/10.2135/cropsci2009.02.0083>

653 Muthu V, Abbai R, Nallathambi J, Rahman H, Ramasamy S, Kambale R, Thulasinathan T,  
654 Ayyenar B, Muthurajan R (2020) Pyramiding QTLs controlling tolerance against  
655 drought, salinity, and submergence in rice through marker assisted breeding. *PLoS*  
656 *ONE* 15:e0227421. <https://doi.org/10.1371/journal.pone.0227421>

657 Panda D, Mishra SS, Behera PK (2021) Drought tolerance in rice: focus on recent mechanisms  
658 and approaches. *Rice Sci* 28 (2):119-132. <https://doi.org/10.1016/j.rsci.2021.01.002>

659 Pandey S, Bhandari H, Ding S, Prapertchob P, Sharan R, Naik D, Taunk SK, Sastri A (2007)  
660 Coping with drought in rice farming in Asia: insights from a cross-country  
661 comparative study. *Agric Econ* 37:213–224. [https://doi.org/10.1111/j.1574-](https://doi.org/10.1111/j.1574-0862.2007.00246.x)  
662 [0862.2007.00246.x](https://doi.org/10.1111/j.1574-0862.2007.00246.x)

663 Peng S, Huang J, Sheehy JE, Laza RC, Romeo M, Visperas, Zhong X, Centeno GS, Khush GS,  
664 Cassman KG (2004) Rice yields decline with higher night temperature from global  
665 warming. *PNAS* 101:9971–9975

666 Philipp N, Weise S, Oppermann M, Börner A, Keilwagen J, Kilian B, Arend D, Zhao Y,  
667 Graner A, Reif JC, Schulthess AW (2019) Historical phenotypic data from seven

668 decades of seed regeneration in a wheat ex situ collection. *Sci Data* 6:137.  
669 <https://doi.org/10.1038/s41597-019-0146-y>

670 Piepho HP, Möhring J (2007) Computing Heritability and Selection Response From  
671 Unbalanced Plant Breeding Trials. *Genetics* 177:1881–1888.  
672 <https://doi.org/10.1534/genetics.107.074229>

673 Piepho HP, Möhring J, Melchinger AE, Büchse A (2008) BLUP for phenotypic selection in  
674 plant breeding and variety testing. *Euphytica* 161:209–228.  
675 <https://doi.org/10.1007/s10681-007-9449-8>

676 Piepho HP, Möhring J, Schulz-Streeck T, Ogutu JO (2012) A stage-wise approach for the  
677 analysis of multi-environment trials. *Biom J* 54:844–860.  
678 <https://doi.org/10.1002/bimj.201100219>

679 R Core Team (2020) R: A language and environment for statistical computing. R foundation  
680 for statistical computing, Vienna, Austria. <https://www.R-project.org/>

681 Sandhu N, Dixit S, Swamy BPM, Raman A, Kumar S, Singh SP, Yadav RB, Singh ON,  
682 Reddy JN, Anandan A, Yadav S, Venkataeshwarllu C, Henry A, Verulkar S, Mandal  
683 NP, Ram T, Badri J, Vikram P, Kumar A (2019) Marker assisted breeding to develop  
684 multiple stress tolerant varieties for flood and drought prone areas. *Rice* 12:8.  
685 <https://doi.org/10.1186/s12284-019-0269-y>

686 Sandhu N, Kumar A (2017) Bridging the rice yield gaps under drought: QTLs, genes, and their  
687 use in breeding programs. *Agronomy* 7:27. <https://doi.org/10.3390/agronomy7020027>

688 Sandhu N, Singh A, Dixit S, Sta Cruz MT, Maturan PC, Jain RK, Kumar A (2014)  
689 Identification and mapping of stable QTL with main and epistasis effect on rice grain  
690 yield under upland drought stress. *BMC Genet* 15:63. [https://doi.org/10.1186/1471-](https://doi.org/10.1186/1471-2156-15-63)  
691 [2156-15-63](https://doi.org/10.1186/1471-2156-15-63)

692 Sandhu N, Yadav S, Catolos M, Sta Cruz MT, Kumar A (2021) Developing Climate-Resilient,  
693 Direct-Seeded, Adapted Multiple-Stress-Tolerant Rice Applying Genomics-Assisted  
694 Breeding. *Front Plant Sci* 12:. <https://doi.org/10.3389/fpls.2021.637488>

695 Shavrukov Y, Kurishbayev A, Jatayev S, Shvidchenko V, Zotova L, Koekemoer F, de Groot S,  
696 Soole K, Langridge P (2017) Early flowering as a drought escape mechanism in  
697 plants: how can it aid wheat production?. *Front Plant Sci* 8:1950.  
698 <https://doi.org/10.3389/fpls.2017.01950>

699 Smith AB, Cullis BR (2018) Plant breeding selection tools built on factor analytic mixed  
700 models for multi-environment trial data. *Euphytica* 214:143.  
701 <https://doi.org/10.1007/s10681-018-2220-5>

702 Sofi P, Ara A, Gull M, Rehman K (2019) Canopy temperature depression as an effective  
703 physiological trait for drought screening, drought - detection and solutions, Gabrijel  
704 Ondrasek, IntechOpen, <https://doi.org/10.5772/intechopen.85966>

705 STRASA Legacy site - Climate-smart rice. <https://strasa.irri.org/>. Accessed 5 Aug 2021

706 STRASA Legacy site - Climate-smart rice. <https://strasa.irri.org/variatal-releases/drought>.  
707 Accessed 5 Aug 2021

708 Van oort PAJ, Zwart SJ (2018) Impacts of climate change on rice production in Africa and  
709 causes of simulated yield changes. *Glob Change Biol.* 24:1029– 1045.  
710 <https://doi.org/10.1111/gcb.13967>

711 Venuprasad R, Dalid CO, Del Valle M, Zhao D, Espiritu M, Sta Cruz MT, Amante M, Kumar  
712 A, Atlin GN (2009) Identification and characterization of large-effect quantitative trait  
713 loci for grain yield under lowland drought stress in rice using bulk-segregant analysis.  
714 *Theor Appl Genet* 120(1):177-90. <https://doi.org/10.1007/s00122-009-1168-1>

715 Vergara GV, Collard BCY, Pamplona A, Gregorio GB (2016). Fast-tracking rice varietal  
716 testing and adoption in different ecosystems in the Philippines. In: 36th Rice Technical  
717 Working Group (RTWG) meeting, Galveston, Texas, US.  
718 <https://doi.org/10.13140/RG.2.1.2575.2727>

719 Vikram P, Swamy BM, Dixit S, Ahmed HU, Sta Cruz MT, Singh AK, Kumar A (2011)  
720 qDTY1.1, a major QTL for rice grain yield under reproductive-stage drought stress  
721 with a consistent effect in multiple elite genetic backgrounds. *BMC Genet* 12:89.  
722 <https://doi.org/10.1186/1471-2156-12-89>

- 723 Wasaya A, Zhang X, Fang Q, Yan Z (2018) Root phenotyping for drought tolerance: a review.  
724 *Agronomy* 8(11):241. <https://doi.org/10.3390/agronomy8110241>
- 725 Wing RA, Purugganan MD, Zhang Q (2018) The rice genome revolution: from an ancient  
726 grain to green super rice. *Nat Rev Genet* 19:505–517. <https://doi.org/10.1038/s41576-018-0024-z>
- 727
- 728 Xu Y, Li P, Zou C, Lu Y, Xie C, Zhang X, Prasanna BM, Olsen MS (2017) Enhancing genetic  
729 gain in the era of molecular breeding. *J Exp Bot* 68:2641–2666.  
730 <https://doi.org/10.1093/jxb/erx135>
- 731 Yadav S, Sandhu N, Dixit S, Singh VK, Catolos M, Mazumder RR, Rahman MA, Kumar  
732 (2021) Genomics-assisted breeding for successful development of multiple-stress-  
733 tolerant, climate-smart rice for southern and southeastern Asia. *Plant Genome*  
734 14:e20074. <https://doi.org/10.1002/tpg2.20074>
- 735 Yadaw RB, Raman A, Kumar A, Swamy BPM, Mishra KK, Sta Cruz MT, Pandey M, Maturan  
736 PT, Vikram P, Dixit S (2013) A QTL for high grain yield under lowland drought in the  
737 background of popular rice variety Sabitri from Nepal. *Field Crops Res* 144:281–287.  
738 <https://doi.org/10.1016/j.fcr.2013.01.019>

### 739 **Acknowledgements**

740 We express profound gratitude to the Bill and Melinda Gates Foundation to fund this research  
741 study. Special thanks to Hans Bhardwaj (Platform Leader, Rice Breeding Innovations, IRRI) for  
742 his guidance and all the support. Authors also thank IRRI rainfed breeding program breeders  
743 Gary Atlin, Arvind Kumar and Shalabh Dixit who all had conducted these trials since 2003 and  
744 generate the data used in this study. We also thanks Jessica Rutkoski a former IRRI biometrician  
745 who initiated the genetic gain study in the IRRI’s drought breeding program and we adopted  
746 some of the R scripts from her work.

### 747 **Funding**

748 The study was supported by the AGGRi Alliance project, “Accelerated Genetic Gains in Rice  
749 Alliance” funded by the Bill and Melinda Gates Foundation.

750 **Author Information**

751 **Affiliations**

752 **Rice Breeding Innovation Platform, International Rice Research Institute (IRRI), Los**  
753 **Banos, Laguna 4031, Philippines**

754 Apurva Khanna, Mahender Anumalla, Margaret Catolos, Jérôme Bartholomé, Roberto Fritsche-  
755 Neto, John Damien Platten, Daniel Joseph Pisano, Alaine Gulles, Ma Teresa Sta. Cruz, Joie  
756 Ramos, Gem Faustino, Sankalp Bhosale, Waseem Hussain

757 **CIRAD, UMR AGAP, Montpellier, France, AGAP, Univ Montpellier, CIRAD, INRA,**  
758 **Montpellier SupAgro, Montpellier, France.**

759 Jérôme Bartholomé

760

761 **Author contributions**

762 WH and SB designed the study, conceived the concept, and provided overall guidance in  
763 conducting the study. AK and WH performed all the statistical analysis. AK wrote the initial  
764 draft of the manuscript. MA and MC helped in data collections, analysis and drafting the  
765 manuscript. JB, RFN and JDP helped in editing and reviewing the manuscript and provided  
766 overall guidance in the data analysis. JB, DJP and AG helped in the data analysis and extraction  
767 of pedigrees for data. TSC, JR and GF helped in data collections and drafting the manuscript. All  
768 the authors read, reviewed, and approved the final manuscript.

769 **Corresponding author**

770 Waseem Hussain

771 International Rice Research Institute (IRRI), Los Baños, Laguna 4031, Philippines,

772 email: [waseem.hussain@irri.org](mailto:waseem.hussain@irri.org)

773

774 **Ethics declarations**

775 **Ethics approval and consent to participate**

776 Not applicable

777 **Competing interests**

778 The authors declare that they have no competing interests.

779 **Publisher's Note**



780 **Supplementary Information**

781 **Genetic Trends Estimation in IRRIs Rice Drought Breeding Program and Identification of**  
782 **High Yielding Drought-tolerant Lines**

783 Apurva Khanna<sup>1</sup>, Mahender Anumalla<sup>1</sup>, Margaret Catolos<sup>1</sup>, Jerome Bartholome<sup>2</sup>, Roberto  
784 Fritsche-Neto<sup>1</sup>, John Damien Platten<sup>1</sup> Daniel Joseph Pisano<sup>1</sup>, Alaine Gulle<sup>1</sup>, Ma Teresa Sta.  
785 Cruz<sup>1</sup>, Joie Ramos<sup>1</sup>, Gem Faustino<sup>1</sup>, Sankalp Bhosale<sup>1</sup>, Waseem Hussain<sup>1\*</sup>

786 <sup>1</sup>Rice Breeding Innovation Platform, International Rice Research Institute (IRRI), Los Banos,  
787 Laguna, Philippines

788 <sup>2</sup>CIRAD, UMR AGAP, Montpellier, France, AGAP, Univ Montpellier, CIRAD, INRA,  
789 Montpellier SupAgro, Montpellier, France

790

791 **\* Corresponding author:**

792 **Waseem Hussain**

793 International Rice Research Institute (IRRI), Los Baños, Laguna 4031, Philippines,  
794 email: [waseem.hussain@irri.org](mailto:waseem.hussain@irri.org)

795

796 **Additional file 1: Figure S1:** Boxplots showing distribution of raw data values for days to 50%  
797 flowering (DTF) data under non-stress and drought conditions.

798 **Additional file 1: Figure S2:** Boxplots showing the distribution of raw data values for plant  
799 height(cm) data under non-stress and drought conditions.

800 **Additional file 1: Figure S3:** Depiction of the number of same genotypes tested across the years  
801 in the breeding program.

802 **Additional file 1: Figure S4:** Pedigree-based heat map and clustering of the genotypes bred over  
803 the years.

804 **Additional file 1: Figure S5:** Distribution of the breeding values for grain yield.

805 **Additional file 1: Figure S6:** Shows genetic trends from the year 2005- 2012 under non-stress  
806 conditions, b) shows genetic trends from the year 2013- 2019 under non-stress conditions, c)  
807 shows genetic trends from the year 2003- 2012 under stress conditions, and d) shows genetic

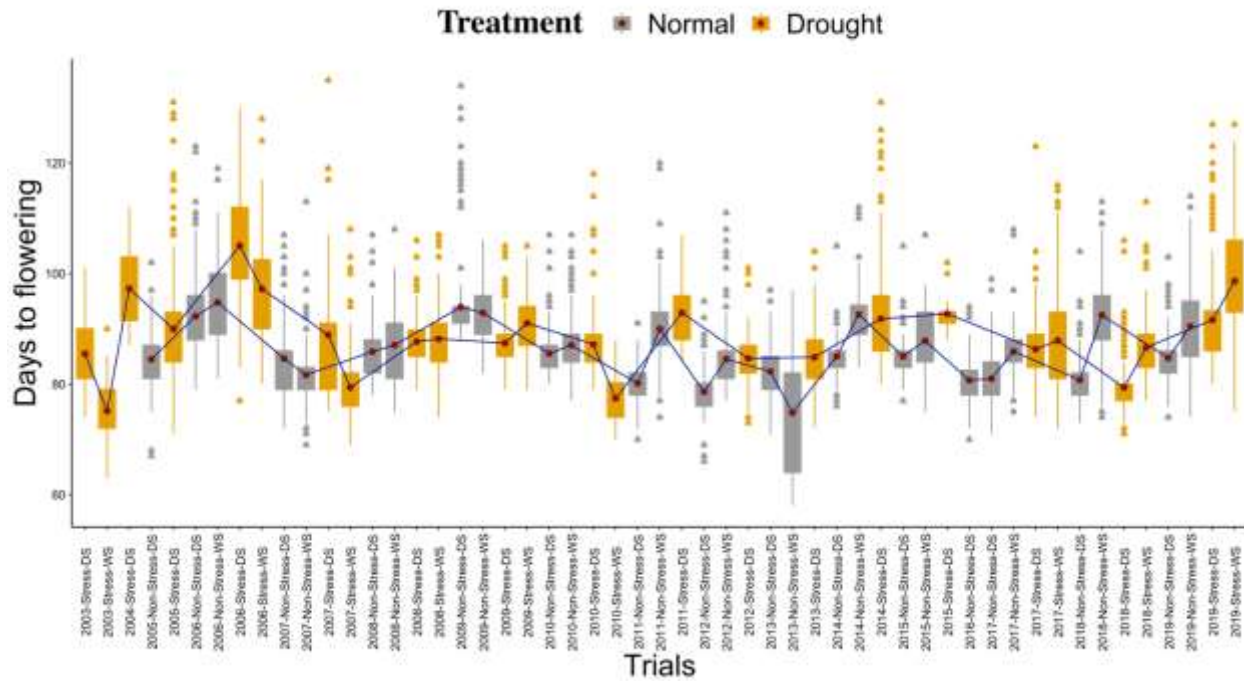
808 trends from the year 2013- 2019 under stress conditions. Both under non-stress and stress  
809 conditions higher genetic gain was observed post to year 2012.

810 **Additional file 1: Figure S7:** Distribution of 200 selected genotypes for grain yield breeding  
811 values.

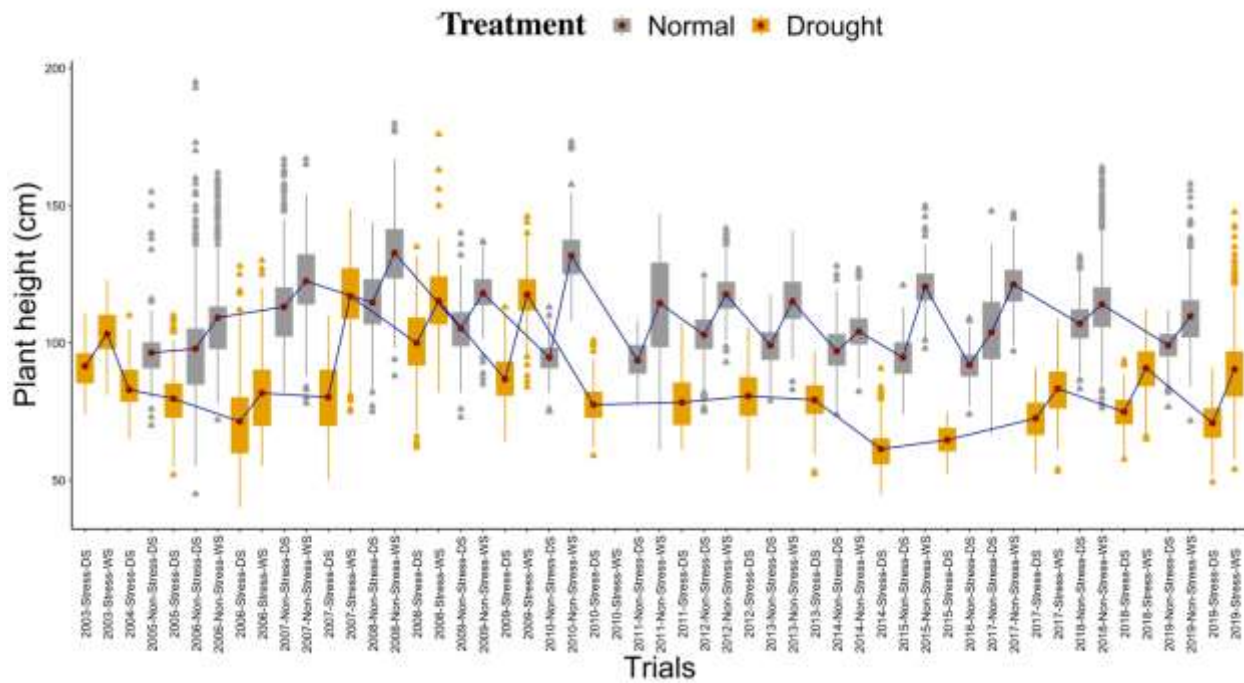
812 **Additional file 1: Figure S8:** Breeding schemes implemented each year in the drought breeding  
813 program from the year 2003-2019.

814 **Additional file 1: Table S1:** List of trials used in this study for genetic trend estimations and  
815 formulation of elite core panel.

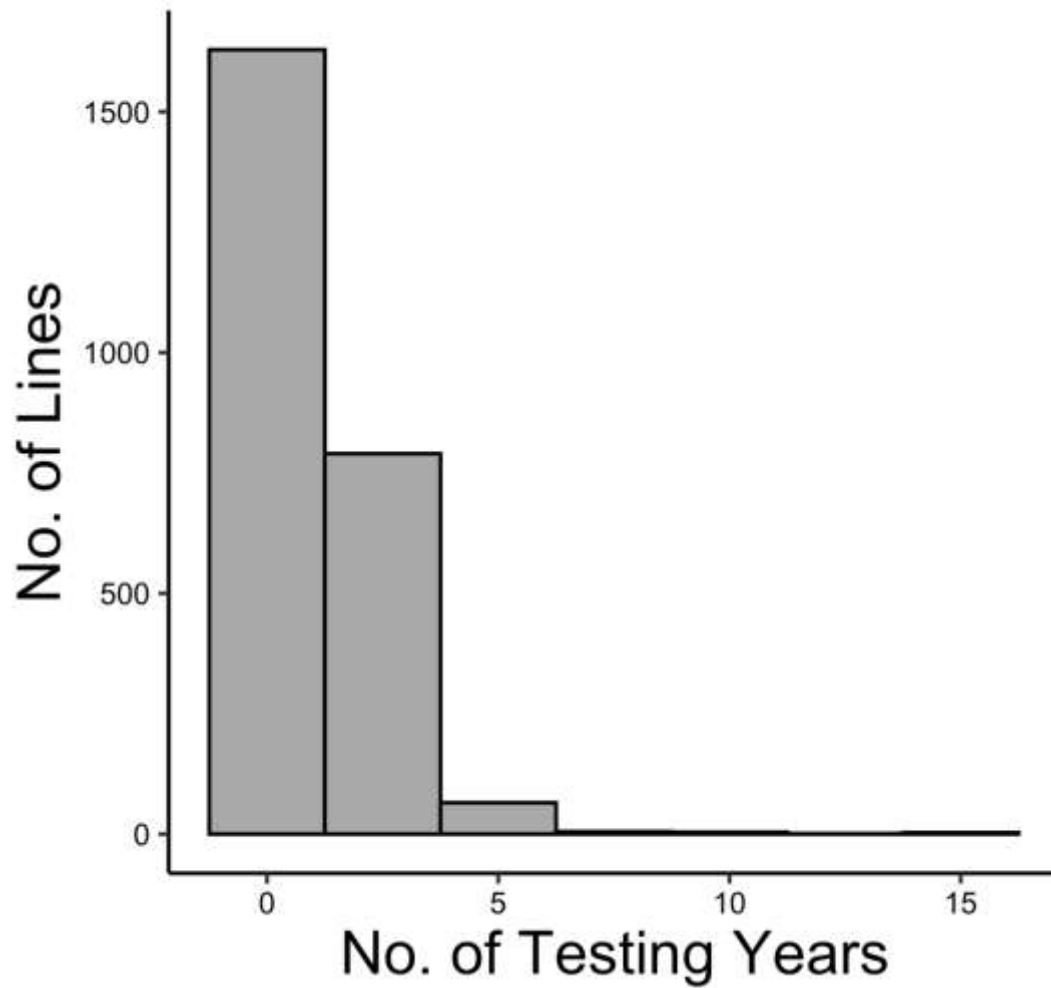
816 **Additional file 1: Table S2:** Delineation of the traits and characteristics of top performing  
817 genotypes formulating the breeding panel.



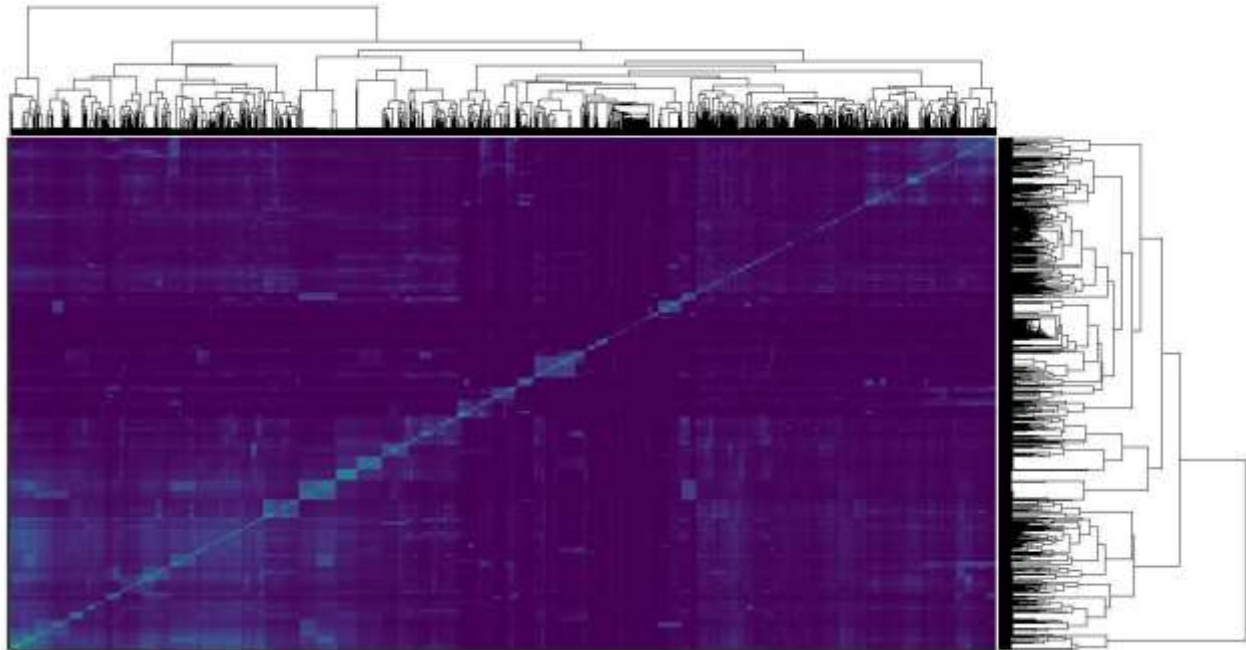
818  
 819 **Figure S1:** Boxplots showing the raw days to 50% flowering (DTF) data under non-stress  
 820 (normal) conditions and drought conditions from the year 2003-2019. The Plant height values of the  
 821 non-stress trials are illustrated in gray, as a comparison to the Plant height values of the stress  
 822 trials denoted in orange.



823  
 824 **Figure S2:** Boxplots showing the raw plant height (cm) data under non-stress (normal)  
 825 conditions and drought conditions from the year 2003-2019. The plant height values of the non-  
 826 stress trials are illustrated in gray, as a comparison to the plant height values of the stress trials  
 827 denoted in orange.

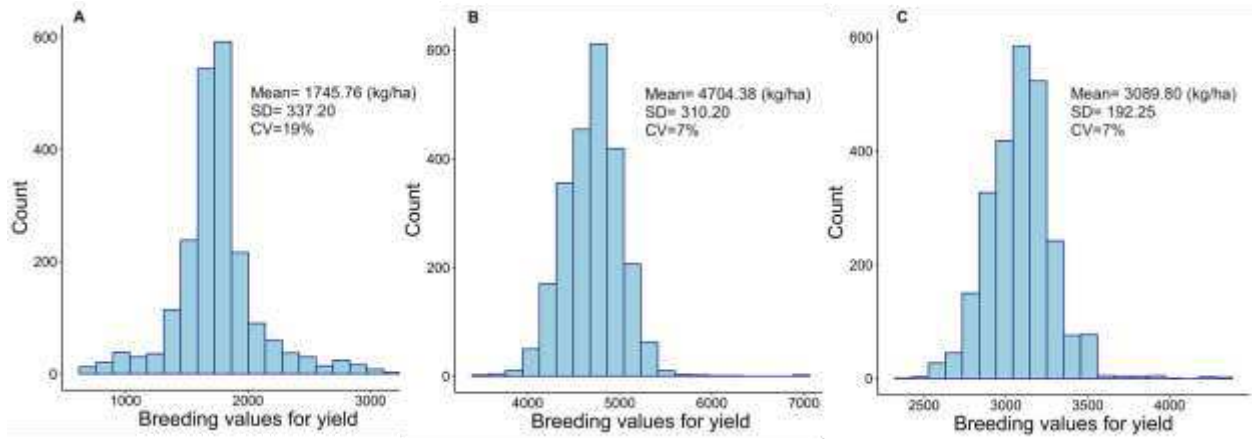


828  
 829 **Figure S3:** The figure depicts the number of the same genotypes tested across the years. It is  
 830 apparent from the figure that a limited number of genotypes were evaluated for more than one  
 831 year. Less than 50 genotypes were evaluated for more than two years.



832  
833

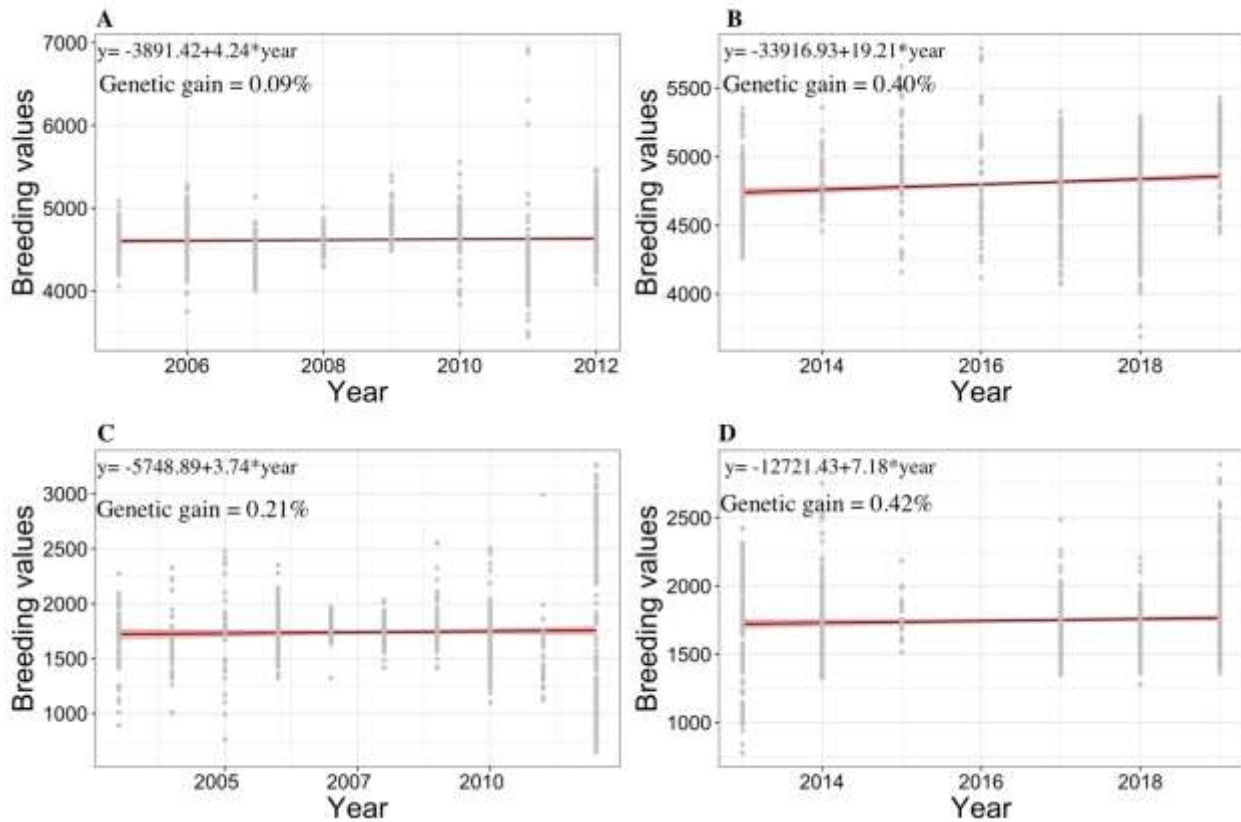
834 **Figure S4:** Heatmap and grouping of pedigree-based relationship matrix based on 2,490 unique  
835 genotypes utilized in IRRI's rice drought breeding program for 17 years from 2003 to 2019. The  
836 figure depicts connectivity between all the lines. The genetically similar lines are depicted in  
837 celeste color which is very well established across the diagonals and evident in the off diagonals  
838 indicating the genetic similarity between different genotypes of the breeding program. Each  
839 square in the figure represents the sub-group and in total 4 main groups were evident indicating  
840 ample diversity in the historical collection.



841  
842  
843  
844

**Figure S5:** Distribution of the breeding values for grain yield (kg/ha) under three conditions: a) Stress, b) Non-Stress, and c). Combined stress and non-stress.

845



846

847

848

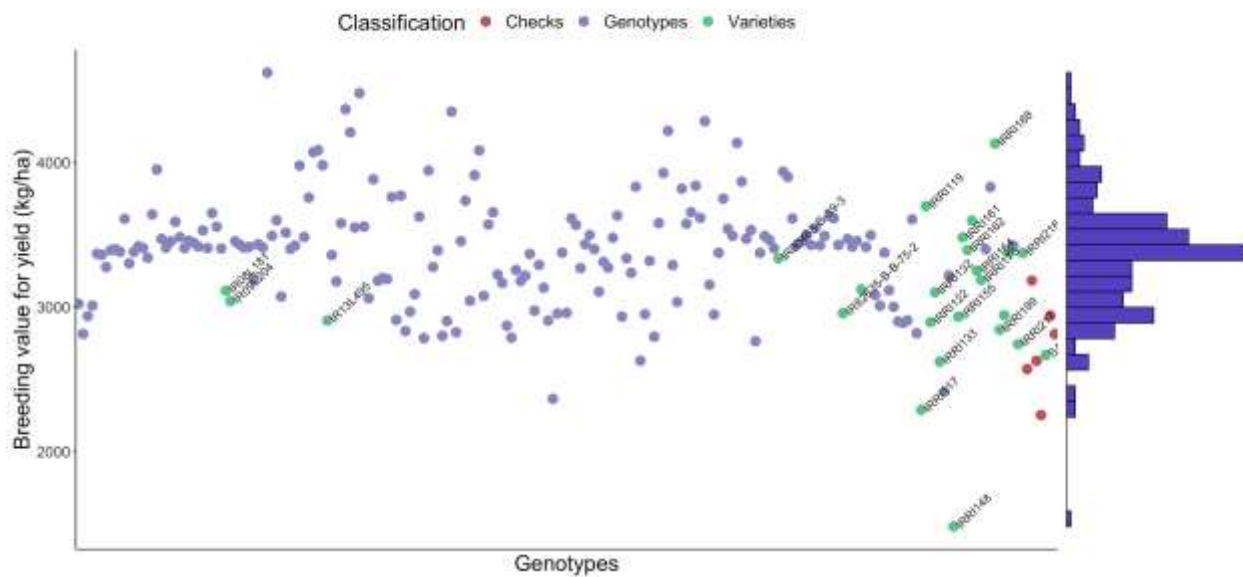
849

850

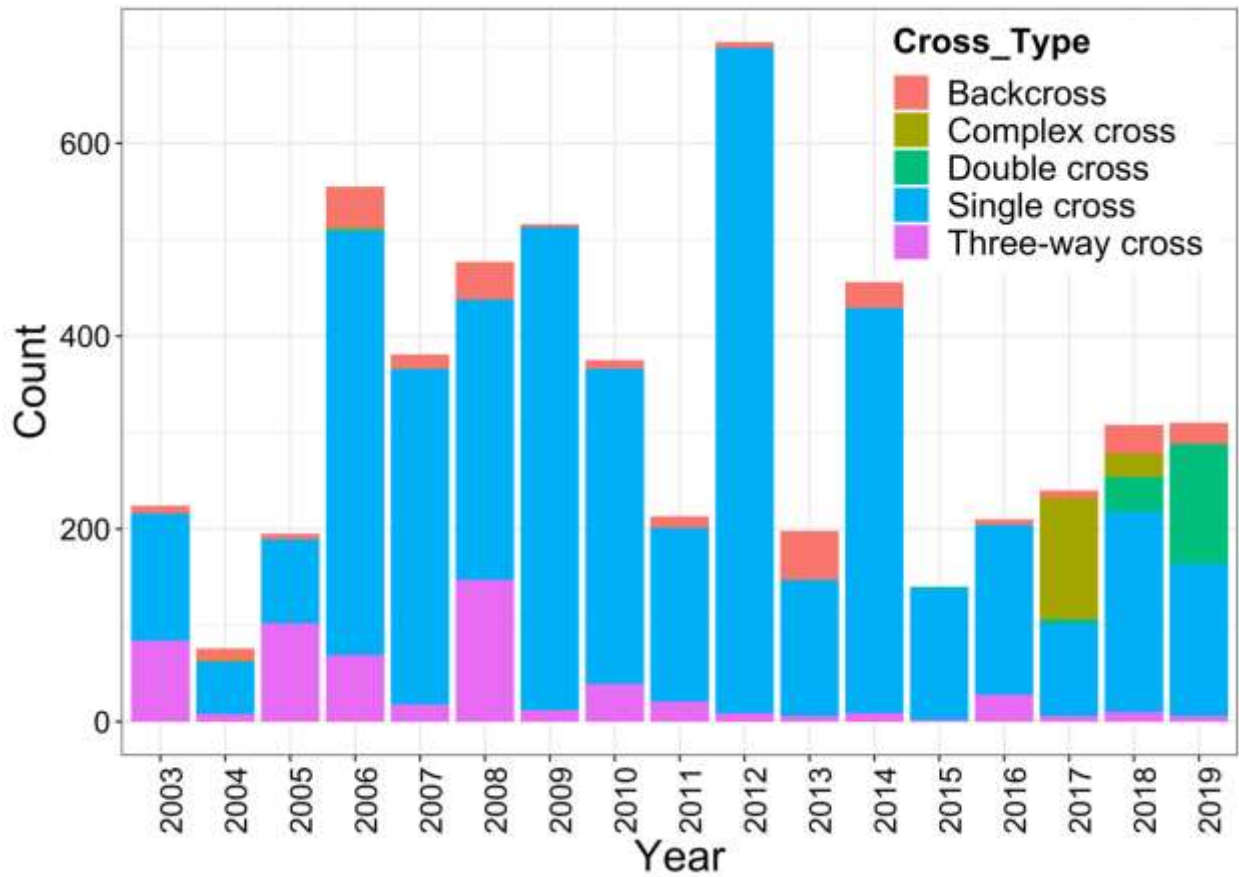
851

**Figure S6:** a) Shows genetic trends from the year 2005- 2012 under non-stress conditions, b) shows genetic trends from the year 2013- 2019 under non-stress conditions, c) shows genetic trends from the year 2003- 2012 under stress conditions, and d) shows genetic trends from the year 2013- 2019 under stress conditions. Both under non-stress and stress conditions higher genetic gain was observed post to year 2012.





852  
 853 **Figure S7:** Representation of the distribution of genotypes comprising the breeding panel. The  
 854 genotypes include superior genotypes with high breeding values and reliabilities along with  
 855 additional traits for tolerance to biotic abiotic stresses with superior grain and cooking quality  
 856 traits. The blue dots represent the selected breeding panel lines which as depicted by the figure  
 857 have higher or comparable breeding values compared to the drought-tolerant checks in red and  
 858 IRRI released varieties in green dots, respectively.



859  
 860 **Figure S8:** The information stating year-wise classification of genotypes developed through  
 861 various breeding strategies from the year 2003-2019. Mixed types of crossing strategies were  
 862 adopted by the breeders. In later years more backcross and complex crosses were made.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.xlsx](#)