Check for updates

ARTICLE

Population improvement via recurrent selection drives genetic gain in upland rice breeding

Adriano Pereira de Castro [1✉], Flávio Breseghello[1], Isabela Volpi Furtini[1], Marley Marico Utumi[2], José Almeida Pereira[3], Tuong-Vi Cao[4,5] and Jérôme Bartholomé [4,5,6]

One of the main challenges of breeding programs is to identify superior genotypes from a large number of candidates. By gradually increasing the frequency of favorable alleles in the breeding population, recurrent selection improves the population mean for target traits, increasing the chance to identify promising genotypes. In rice, population improvement through recurrent selection has been used very little to date, except in Latin America. At Embrapa (Brazilian Agricultural Research Corporation), the upland rice breeding program is conducted in two phases: population improvement followed by product development. In this study, the CNA6 population, evaluated over five cycles (3 to 7) of selection, including 20 field trials, was used to assess the realized genetic gain. A high rate of genetic gain was observed for grain yield, at 215 kg.ha^{-1} per cycle or 67.8 kg.ha^{-1} per year (3.08%). The CNA6 population outperformed the controls only for the last cycle, with a yield difference of 1128 kg.ha^{-1}. An analysis of the product development pipeline, based on 29 advanced yield trials with lines derived from cycles 3 to 6, showed that lines derived from the CNA6 population had high grain yield, but did not outperform the controls. These results demonstrate that the application of recurrent selection to a breeding population with sufficient genetic variability can result in significant genetic gains for quantitative traits, such as grain yield. The integration of this strategy into a two-phase breeding program also makes it possible to increase quantitative traits while selecting for other traits of interest.

INTRODUCTION

In the current context of population growth, with estimates of a world population of 9.8 billion people in 2050 and 11.2 billion in 2100 (FAO 2017) requiring more efficient natural resource use, rice production in Latin America and Sub-Saharan Africa has a considerable potential for expansion. In this scenario of increasing demand and climate change, upland rice cultivation could play an important role, particularly given the land and water limitations on the expansion of irrigated rice. Embrapa (Brazilian Agricultural Research Corporation) is responsible for the national rice breeding program in Brazil, which aims to increase the competitiveness, resilience, and profitability of rice in tropical ecosystems, including uplands. This breeding program is based on a strategy of achieving genetic gains while preserving genetic variability to meet long-term objectives of improving the traits of interest (Barros et al. 2018; Martínez et al. 2014). A two-phase breeding scheme is used, consisting of a population improvement phase, in which the aim is to improve performance for quantitative traits, and a product development phase focusing on the fixation and selection of elite material (Fig. 1). The upland rice breeding program at Embrapa is based on four populations developed for high yield with medium cycle, high yield with short cycle, drought tolerance and low phosphorous tolerance. These populations have been systematically improved, with breeding efforts focusing

mainly on productivity and grain quality but also on tolerance to water deficiency, disease resistance, lodging, greater precocity, and initial vigor.

Most traits of economic importance are quantitative, controlled by many genes and greatly influenced by the environment (Walsh and Lynch, 2009). In this context, recurrent selection (RS) is one of the most efficient breeding methods available. It aims to increase the probability of finding superior genotypes by progressively concentrating favorable alleles and increasing the mean value of the population at each cycle (Dudley 2007; Hallauer and Darrah 1985; Morais 2005). RS strategies generally begin with the constitution of a population with a broad genetic base, followed by multiple cycles of selection and recombination to achieve continuous genetic gain, while minimizing the loss of genetic variability (Rutkoski, 2019). In the selection of superior genotypes for quantitative traits governed by a large number of loci with small effects, the probability of finding individuals combining all the favorable alleles is very small. The selection and intercrossing of superior individuals within a genetically diverse population increase the frequency of favorable alleles in the offspring, thereby increasing the chances of finding families with a better performance. As a result, the population mean for target traits gradually shifts in the desired direction (Hallauer et al. 2010; Morais Júnior et al., 2015). This is the bases of RS, which is now
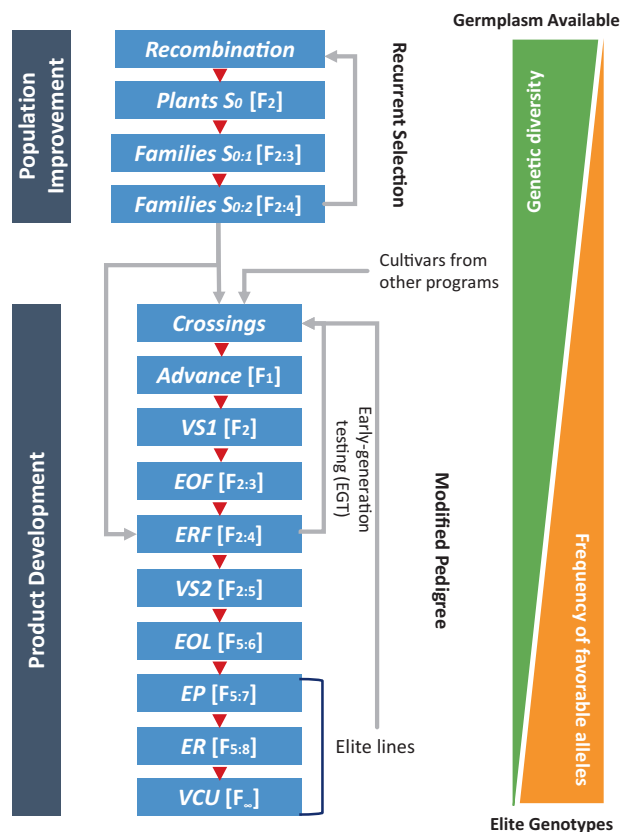
Fig. 1 Schematic representation of the Embrapa rice breeding program, including the population improvement phase and the product development phase. VS1: plant selection nursery 1; EOF: family observation trial; ERF: multisite family yield trial; VS2: plant selection nursery 2; EOL: line observation trial; EP: multisite preliminary trials; ER: multisite regional trials; VCU: multisite trials assessing value for cultivation and use.

considered the best strategy for achieving genetic gains for quantitative traits in plant species (Bernardo, 2020; Cobb et al., 2019; Grenier et al., 2015).

Plant breeding demands long-term investment. It is, therefore, important to monitor the effectiveness of breeding strategies continually, through the use of indicators such as genetic gain (Ceccarelli, 2015). Genetic gain is the increase in mean breeding value of a population over at least one cycle of selection for a particular trait or index of traits (Eberhart, 1964; Rutkoski, 2019). Genetic gain can be estimated by a linear regression analysis of mean genetic performance over time. In crop species, realized genetic gain is typically estimated from two different types of information: (i) studies evaluating lines or varieties from different time periods in the same trials, possibly in multiple locations and/ or years (known as era studies) or (ii) historical data from a breeding program or from the variety release system (Laidig et al., 2014; Rutkoski, 2019). Phenotypic data stored over time and available for several cycles of selection can be used to estimate the changes associated with past selection, which can guide the orientation of subsequent breeding cycles (Breseghello et al., 2021). The systematic assessment of genetic gains for key traits is, therefore, essential for assessing the efficiency of the breeding scheme (Rutkoski, 2019).

One of the first rice breeding programs based on population improvement through RS began in 1989, through a partnership between Embrapa and CIRAD (*Centre de Coopération Internationale en Recherche Agronomique pour le Développement*) (Taillebois and Guimaraes, 1989). That initiative had the objective of

developing populations with a broad genetic base and a large potential for the generation of promising families. The RS method is used widely in outcrossing species, but much less so in self-pollinating species, due to the difficulty of making a large number of crosses. However, the discovery of recessive genetic male sterility in rice (Fujimaki, 1979) rendered the RS method feasible for this species. The first upland rice population developed at Embrapa was named CNA-IRAT 5, which had a focus on water deficiency tolerance and blast resistance (Taillebois and Guimaraes, 1989). CNA6 was generated in 1993/94 by incorporating the alleles from 27 parents into the CNA-IRAT 5, with the objective of increasing grain yield and drought tolerance.

In this study, our aim was to assess the effectiveness of population improvement via RS for achieving genetic progress, maintaining genetic variability, and increasing the potential for the selection of superior lines. We analyzed data from the five last cycles of the CNA6 population. Genetic parameters for plant height, days to flowering and grain yield were calculated from the data of multi-environment trials performed in the rice-growing regions of Brazil. The contribution of the CNA6 population to the development of elite lines was assessed. Additionally, we suggest adjustments to the RS methodology, with the aim of increasing the efficiency of family selection in the upland rice breeding program.

MATERIALS AND METHODS
Upland rice breeding program structure
The upland rice breeding program at Embrapa has two phases (Fig. 1). The aim of the first phase is to improve the base population through RS. The second phase focuses on product development through pedigree breeding using elite material selected from the improved population and from other breeding programs.

Population improvement. The CNA6 population was created in 1993/94 by crossing 27 parents with the CNA-IRAT 5 population. The CNA-IRAT 5 population was developed in 1984 by crossing 26 *japonica* lines with the male-sterile mutant of IR 36 (*indica*), developed by Singh and Ikehashi (1981). This mutant carries the recessive *ms* male sterility gene, facilitating the recombination between families (Fujimaki, 1979).

The first two breeding cycles for the CNA6 population were based on mass selection in the S_0 generation (Table S2). This selection was based mainly on flowering time and plant height, with the aim of eliminating late-maturing and/or tall genotypes. In each cycle we selected about 100 plants to limit the effects of genetic drift. In the second cycle, we carried out a field evaluation of yield (YLD) based on $S_{0:1}$ families. The remaining seeds from the selected families were combined to generate the population for the third cycle. This population was then subjected to five cycles of selection, as follows: (i) selection of individual male-sterile plants in the S_0 generation; (ii) evaluation of $S_{0:1}$ families and selection of the top 30% by bulk-harvested, retaining five panicles from each family (each panicle corresponds to a $S_{1:2}$ progeny) as recombination units; (iii) evaluation of $S_{0:2}$ families in multisite yield trials; (iv) $S_{1:2}$ progenies from the best families were recombined. Recombination was performed with genetic male sterility (*msms* genotype) in cycles 3, 4 and 5, and by manual crosses in cycles 6 and 7. The duration of the cycle was three years for cycles 3 to 6 and four years for cycle 7. Therefore, the average cycle duration over the five cycles was 3.2 years.

Individual plants from the S_0 generation were selected on the basis of plant architecture, disease resistance and grain type (long-slender), according to the requirement of the Brazilian consumer market. Selection in $S_{0:1}$ families focused on high-heritability traits, including plant height, days to flowering, lodging resistance, disease resistance and grain quality. Selection of $S_{0:2}$ families focused on grain yield. These target traits for the upland rice breeding program in Brazil have been described in more detail elsewhere (Breseghello et al. 2011; Colombari Filho et al. 2013).

Product development. Embrapa's product development phase follows a modified pedigree method, with early-generation testing (EGT) as a strategy for selecting segregating families with potential for extraction of high-yielding inbred lines (Barros et al., 2018). The EGT approach is based on the premise that the performance of segregating families in early

Table 1. Characteristics of the yield trials of $S_{0:2}$ families for each cycle of recurrent selection on the CNA6 population.

Cycle	Year	Sites[a]	Number of $S_{0:2}$ families	Experimental design	Controls[b]
3	2000/01	SAG, PRI, VIL	220	Augmented block	C1, C2, C3, C4
4	2003/04	SAG, GOI, SIN, LAV, TER	92	Augmented block	C1, C2, C3, C5
5	2006/07	SAG, SIN, TER	200	Augmented block	C1, C2, C3, C5
6	2009/10	SAG, POR, SIN, PAR	61	Augmented block	C1, C5, C6, C7
7	2014/15	SAG, POR, SIN	93	Lattice	C1, C2, C3, C5, C6, C7, C8

[a]SAG: Santo Antônio de Goiás, GO; PRI: Primavera do Leste, MT; VIL: Vilhena, RO; GOI: Goianira, GO; SIN: Sinop, MT; LAV: Lavras, MG; TER: Teresina, PI; POR: Porangatu, GO; PAR: Paragominas, PA.
[b]C1: BRS Primavera; C2: Carajás; C3: Guarani; C4: CNA8097; C5: BRS Bonança; C6: BRSMG Curinga; C7: BRS Sertaneja; C8: BRS Esmeralda.

Table 2. Characteristics of the advanced yield trials including elite lines derived from recurrent selection on the CNA6 population in the product development pipeline. EP, ER and VCU are the preliminary yield trials, regional yield trials and trials assessing value for cultivation and use, respectively.

Trial	Season	Design	Trials	Replicates	Group		
					CNA6 Lines	Other Elite Lines	Controls
EP	2016/17	Lattice 6 ×6	4	3	34	0	2
EP	2017/18	Lattice 9 ×9	4	2	10	69	2
ER	2017/18	Alfa-lattice 9 ×5	4	3	15	26	4
ER	2018/19	Alfa-lattice 8 ×7	5	3	21	31	4
VCU	2020/21	Randomized blocks	12	4	4	12	4

generations is a good predictor of the mean performance of the inbred lines derived from them (Bernardo, 2003). EGT improves quantitative traits very effectively in elite populations, through the use of selected families with a high additive genetic value as new parents, as in RS, or the selection of inbred lines. This approach combines the evaluation of yield in segregating families and recombination of the best families with elite inbred lines and commercial cultivars, before the advanced inbreeding stage (Barros et al., 2018; van Ginkel and Ortiz 2018). Once inbred lines are obtained, they are evaluated in multisite trials with the objective of selecting candidates for commercial release (Fig. 1). Phenotypic data of the elite lines developed from CNA6 families from multisite trials were used for assessing the potential of CNA6 for the development of upland rice cultivars.

Field experiments and phenotypic traits

Recurrent selection: progeny testing. The experimental data used in this study was generated in multisite trials conducted on $S_{0:2}$ families from five cycles of RS in the CNA6 population (Table 1). The following traits were analyzed: grain yield (YLD, kg.ha^{-1}), days to flowering (FLW, days) and plant height (PHT, cm). In total, 682 $S_{0:2}$ families were evaluated in 20 experiments, conducted in the seasons 2000/01 2003/04 2006/07 2009/10 and 2014/15, corresponding to cycles 3, 4, 5, 6 and 7, respectively. Within each cycle, the same set of families was evaluated at all sites. The characteristics of the 20 field experiments are presented in Tables 1, and S1. The geographic area covered by this study encompasses the states with the highest levels of upland rice production in Brazil (Ramirez-Villegas et al., 2018).

In cycles 3 to 6, trials were conducted with an augmented block design (Federer and Raghavarao, 1975). The trials for the seventh cycle had an alpha-lattice design with two replicates per site. Plots consisted of four rows of 5 meters, with an inter-row spacing of 0.30 meters and 60 seeds per meter sow density corresponding to 1200 plants per plot. All trials were rainfed, conducted in aerobic soil, in the wet season. Chemical fertilizers were applied to the soil at sowing, at rates considered intermediate-to-high relative to standard farm management. Nitrogen fertilizer was applied as a top dressing at tillering stage. Weed control was done by herbicides, supplemented by hand-weeding when needed. Insecticides were applied to the furrow in most trials, with complementary canopy spraying if required. No fungicides were applied, as one of the features evaluated in the trials was the genetic resistance to fungal diseases.

For YLD evaluations, the two central rows were harvested, discarding

0.50 m on each end of the plot. A total of 480 plants per plot (2 rows of 4 meters, with 60 seeds per meter) was used to evaluate grain yield. For PHT evaluations, a mean of five readings for the distance from the base of the plant to the tip of the panicle was obtained. For FLW, the interval between sowing and heading of 50% of the plants was determined.

Selection of families. The best $S_{0:2}$ families in each cycle were selected for recombination and the generation of a new breeding cycle. Data for the selected families were used to estimate the gain associated with RS. The selection differential (S) was calculated as the difference between the mean for the selected families and the population mean. The response to selection, or selection gain, was calculated by multiplying the selection differential by the heritability of the specific trait ($R = h^2S$).

Product development: evaluation of advanced lines. Fifty-six inbred lines were derived from CNA6 $S_{0:2}$ families selected in cycles 3 to 6 and evaluated in multisite yield trials. Given the time required for inbreeding and evaluation, no data were available for lines derived from cycle 7. The lines derived from the CNA6 population were tested in preliminary trials (EP, F_7), regional trials (ER, F_8) and multisite VCU (value for cultivation and use, F_∞) trials (Fig. 1) in the 2016/17 to 2020/21 cropping seasons (Table 2). Some of the best performing lines were evaluated at several stages. The experimental design used was an alpha-lattice with two or three repetitions for EP and ER trials and randomized complete blocks with four repetitions for the VCU trials. The plots and the protocol for the trials were as described for progeny testing above. The traits grain yield (YLD, kg.ha^{-1}), days to flowering (FLW, days) and plant height (PHT, cm), were evaluated. BRS Sertaneja, a control present in cycles 6 and 7 of CNA6, and BRS Esmeralda, present in cycle 7, were also used as controls in the multisite trials of the product development phase.

Statistical analyses

Genetic trend evaluation on the CNA6 population. Statistical analyses were performed to check the quality of each of the 20 trials. We checked the data for outliers with a model-based method deployed within the SAS Proc Mixed procedure, using the "influence" option of the "model" statement. Proc Mixed computes parameter estimates from the difference between full-data and reduced-data estimates. Based on restricted likelihood distances, Cook's distance and CovRatio, 0 to 9 plots per trial were identified as outliers and were removed for subsequent analyses. The outliers accounted for less than 3.5% of the total number of plots. For each
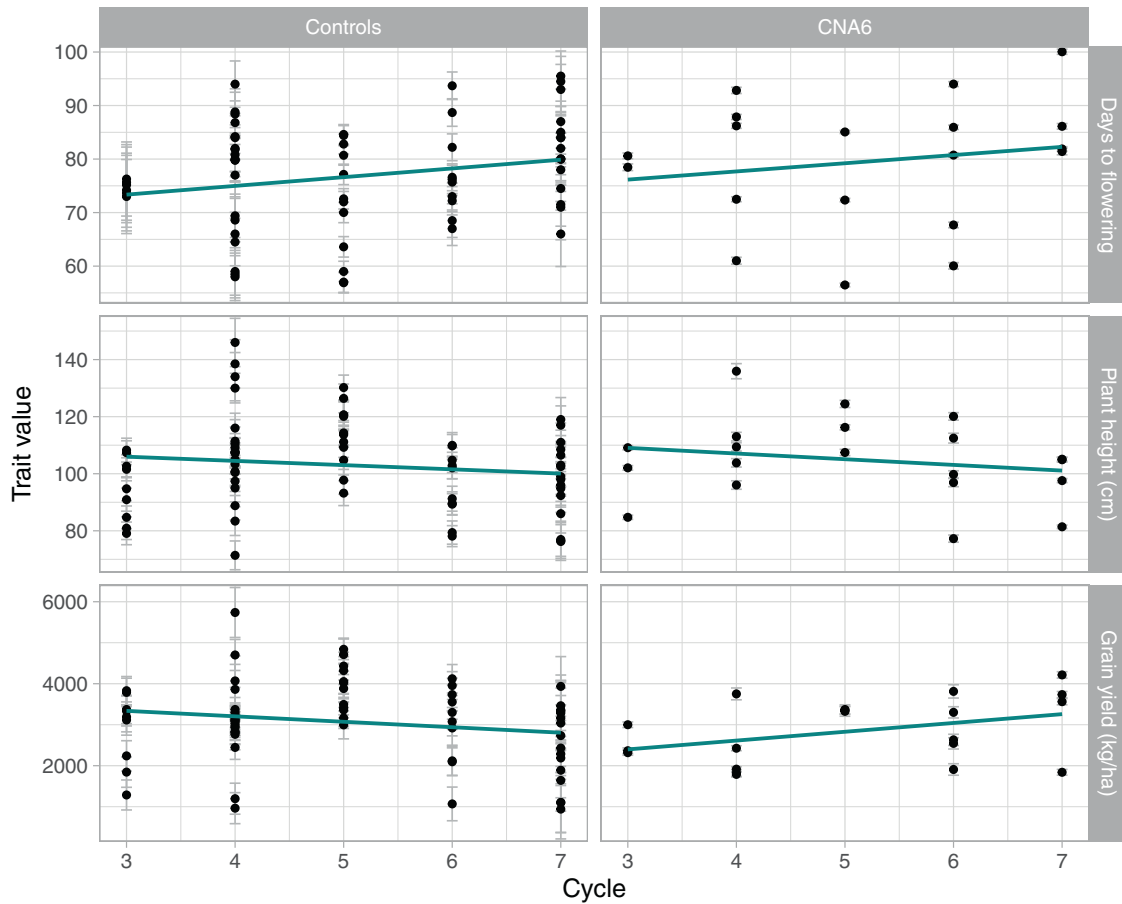
Fig. 2 Trend analysis of CNA6 families against the controls in cycles 3 to 7. Each individual data point corresponds to the least squares mean of the controls or the CNA6 families in one trial (20 trials in total). The blue lines represent the linear regression over the cycles.

cycle and each trait, the Pearson correlation coefficients between trials were estimated using adjusted values at the trial level.

Breeding cycles involved different generations of families, and the same set of families was evaluated across all trials, within a given cycle. We performed a two-step analysis: (1) at plot data level and (2) at genotypic mean level. The first step involved analyzing the plot data to extract the variance components and adjusted means of families. Two models were used, depending on the experimental design:

Augmented blocks design (cycles 3, 4, 5 and 6):

$$Y_{ijkl} = \mu + T_l + C_k + TC_{kl} + b(T)_{jl} + g(C)_{ik} + e_{ijkl}$$

Y_{ijkl}: Observed phenotype
μ: Grand mean
T_l: Fixed effect of trial l from a given cycle, $l \in \{1, 2, ..., L\}$
C_k: Fixed effect of group k, k being either the group of controls or the group of CNA6 families
TC_{kl}: Interaction between trial l and group k
$b(T)_{jl}$: Random block effect j nested in trial l, $b(T)_{jl} \sim N(0, \sigma_b^2)$
$g(C)_{ik}$: Random family or control effect i, nested in group k, $g(C)_{ik} \sim N(0, \sigma_g^2)$,
e_{ijkl}: Random residual, structured by trial, $e_{ijkl} \sim N(0, \sigma_e^2)$, $\sigma_e^2 = (\sigma_{e1}^2, \sigma_{e2}^2, ..., \sigma_{eL}^2)$
Alpha-lattice design (cycle 7):

$$Y_{ijkrl} = \mu + T_l + R(T)_{rl} + C_k + TC_{kl} + b(RT)_{jrl} + g(C)_{ik} + e_{ijkrl}$$

Y_{ijkrl}: Observed phenotype
μ: Grand mean
T_l: Fixed effect of trial l from a given cycle, $l \in \{1, 2, ..., L\}$
$R(T)_{rl}$: Fixed effect of replication r nested in trial l, $r \in \{1, 2\}$
C_k: Fixed effect of group k, k being either the group of controls or the group of CNA6 families

TC_{kl}: Interaction between trial l and group k
$b(RT)_{jrl}$: Random block effect j nested in replication r and trial l, $b(RT)_{jrl} \sim N(0, \sigma_b^2)$
$g(C)_{ik}$: Random family or control effect i, nested in group k, $g(C)_{ik} \sim N(0, \sigma_g^2)$
e_{ijkrl}: Random residual, structured by trial, $e_{ijkrl} \sim N(0, \sigma_e^2)$, $\sigma_e^2 = (\sigma_{e1}^2, \sigma_{e2}^2, ..., \sigma_{eL}^2)$

The adjusted means of the progenies by cycle were used to estimate the genotypic correlation between traits. For each cycle, the Pearson correlation coefficients were computed using SAS. The variance components extracted from the models described above were used to compute the heritability for a given cycle. The following formula was used: $H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$, σ_g^2 being the variance associated with the genetic effect and σ_e^2 the variance associated with the residuals, as explained above.

The second step involved combining the genotypic adjusted means of all cycles together and modeling them as a result of the selection process. We studied the trend in this process, by considering a linear regression model with the two groups (CNA6 families and controls) to estimate the intercept and the slope associated with each group. The model can be written as follows: $\hat{Y}_{it} = a_i + \beta_i \overline{Y}_t + e_{it}$, where a_i is the intercept for the group of controls or the CNA6 population, β_i is the corresponding slope, $\overline{Y}_t$ is the estimated mean for cycle t and e_{it} is the residual term. In the case of the CNA6 population, the slope β_i can be interpreted as the genetic trend across cycles. It corresponds to the mean unit gain per breeding cycle. In the case of the controls, β_i correspond to the non-genetic trend mostly related to environmental variations.

Adjusted performances of inbred lines. The inbred lines were evaluated in multisite trials, in different generations and years (Table 2). Single-trial analyses were performed to evaluate the quality of data. Joint analysis was

performed for multisite trials per year, using the model below:

$$Y_{ijrl} = \mu + T_l + R(T)_{rl} + b(RT)_{jrl} + g_i + gT_{il} + e_{ijrl}$$

Y_{ijrl}: Observed phenotype
μ: Grand mean
T_l: Fixed effect of trial, $l \in \{1, 2, ..., L\}$
$R(T)_{rl}$: Fixed effect of replication r nested in trial l, $r \in \{1, 2, 3\}$
$b(RT)_{jrl}$: Random block effect j nested in replication r and trial l, $b(RT)_{jrl} \sim N(0, \sigma_b^2)$
g_i: Fixed effect of genotype, $i \in \{1, 2, ..., I\}$
gT_{il}: Random interaction between genotype i and trial l, $gT_{il} \sim N\left(0, \sigma_{gT}^2\right)$
e_{ijrl}: Random residual, $e_{ijrl} \sim N(0, \sigma_e^2)$, $\sigma_e^2 = (\sigma_{e1}^2, \sigma_{e2}^2, ..., \sigma_{eL}^2)$

For the joint analysis, we used a mixed model approach based on the MIXED procedure (generalized linear model) in SAS (SAS Institute 1985) to obtain the vector of genotype means, adjusted for the trial effect, for the lines and controls evaluated.

RESULTS

Genetic gain in five cycles of recurrent selection on the CNA6 population

The genetic trends for YLD, PHT and FLW from cycle 3 to 7 are shown in Fig. 2. A significant gain was observed for YLD. The mean gain per cycle was 215 kg.ha^{-1} (Table 3), corresponding to a gain of 9.85% per cycle. Over the five selection cycles, the average cycle length was 3.2 years. This resulted in a gain per year of 67.8 kg.ha^{-1} (3.08%). The trend observed for the controls was negative but not significant due to a large variability between controls (Fig. 2 and Table 3). In the first cycle (cycle 3), the yield for the group of CNA6 families was 2318 kg.ha^{-1}, 24% lower than the mean of the controls, at 3036 kg.ha^{-1}. That was expected, as cycle 3 was the first in which multisite yield trials were performed. However, the relative performances of the families, driven by the selection pressure for YLD, increased steadily over time, outperforming the controls in cycle 7 (+37%, Table 4). In cycle 7, the S$_{0:2}$ families had a mean yield of 4167 kg.ha^{-1}, whereas the mean yield for the controls was 3038 kg.ha^{-1}. A medium selection intensity was applied, with 47 (23%), 45 (47%), 46 (21%), 25 (38%) and 38 (41%) families selected in cycles 3, 4, 5, 6 and 7, respectively. This selection intensity was applied with the aim of promoting genetic gain, whilst simultaneously preserving the genetic variability of the population for subsequent cycles. The selection differential ranged from 84.2 kg.ha^{-1} (cycle 5) to 250.6 kg.ha^{-1} (cycle 7), with a mean value of 152.1 kg.ha^{-1} per cycle (Table 4). However, the low heritabilities compromised the gains from selection. The gains ranged from 4.8 (cycle 5) to 146.3 kg.ha^{-1} (cycle 7, Table 4). Cycle 5 had the lowest heritability (0.06) and cycle 7 the highest (0.41). This difference in heritability was related to the level of correlation between trials within each cycle. In cycles 3 to 6, the correlations between trials were mostly not significant, with values ranging from −0.18 to 0.37 (Table S3).

Table 3. Parameters (intercept and slope) for the linear regression estimated for the CNA6 population and controls.

Trait	Group	Intercept	Slope	Std error	P
FLW	CNA6	74.6	1.53	1.99	0.959
	Controls	71.7	1.63	0.94	0.086
PHT	CNA6	111.1	−1.99	2.84	0.857
	Controls	107.5	−1.48	1.35	0.276
YLD	CNA6	2183.0	215	174.3	0.049
	Controls	3470.9	−133.1	82.8	0.112

The parameters are given for three traits: days to flowering (FLW), plant height (PHT) and grain yield (YLD). The p-values and standard error are shown.

In cycle 7, the adoption of an alpha-lattice design increased the accuracy of the evaluation, and the correlations were all significant (0.3 to 0.65). However, for all the cycles, the effect of genotype-by-environment interactions was significant (Table S4).

The results for FLW and PHT after five cycles of selection revealed a non-significant trend (Table 3, Fig. 2). The estimates of the slope were similar than those of the controls indicating that most of the change was associated with non-genetic factors. In the case of FLW an increase of 1.53 days per breeding cycle, or 0.48 days per year was observed (Table 3). This can partly be explained by the negative genetic correlation between FLW and YLD (−0.2, Table 5). In addition, in the last breeding cycle, the mean FLW for the selected families (82.2 days) was 3 days greater than that for the controls (Table 4). For FLW, phenotyping accuracy was high, with heritabilities ranging from 0.47 to 0.73 (Table 5) and significant correlations between trials for all cycles except cycle 6 (Table S3). In the case of PHT, a decrease of 1.99 cm per cycle was estimated (Table 3, Fig. 2). Only two of the five cycles of selection had a negative selection differential, with a lower PHT for the selected families (Table 4). This was probably related to positive genotypic correlations found between PHT and YLD (Table 5). The heritabilities were lower than for FLW and ranged from 0.15 to 0.40.

Elite lines derived from the CNA6 population

In total, 56 elite lines (at least at the F5:7 generation) developed in the upland rice breeding program were derived from cycles 3 to 6 of the CNA6 population. In the most advanced trial in the product development pipeline (VCU) for the 2020/21 cropping season, 25% of the lines were derived from CNA6 (Table 2). The YLD performances of the lines derived from CNA6 population were generally similar to those of other lines evaluated at the same stage (Fig. 3). For PHT and FLW, performances were more variable between stages, due to environmental effects (Fig. 3). However, the CNA6-derived lines had a slightly shorter cycle. Several CNA6-derived lines outperformed the best controls (especially BRS Esmeralda) at the EP and ER stages: 12 at EP 16/17, 5 at EP 17/18, 2 at ER 17/18 and 7 at ER 18/19. However, for the VCU 20–21 stage, none of the CNA6-derived lines outperformed the best controls (BRS Esmeralda and BRS A502).

DISCUSSION

Genetic progress enabled by recurrent selection

Grain yield. Regular assessment of the genetic gain in breeding programs is a key component of any breeding strategy (Ceccarelli, 2015). Indeed, such assessments can identify the strengths and weaknesses contributing to the realized rate of genetic gain and can guide future breeding activities. Here, the genetic gain for YLD was estimated at 215 kg.ha^{-1} per cycle (67.8 kg.ha^{-1}.year^{-1}) or 9.85% per cycle (3.08% per year). These values are in the same range than the ones that reported in other studies of RS in rice in terms of absolute value (kg.ha^{-1}), but the relative value (%) is similar or higher. For example, in three cycles of selection based on the evaluation of S$_{0:2}$ families, Rangel et al. (2002) estimated a genetic gain for YLD of 259.9 kg.ha^{-1} per cycle, or 4.67% in lowland rice. Similar values were reported by Breseghello et al. (2009) after three cycles of RS on an upland population. They estimated an annual genetic gain of 103.4 kg.ha^{-1} (3.6%), with a better mean performance of the selected population than of the controls in the last cycle. More recently, in a study of the CNA12S population for irrigated rice, Morais Júnior et al. (2017) found a rate of genetic gain for yield of 1.98% per year. Our estimate of genetic gain for YLD was obtained over a longer period, with five cycles of selection. This long-term evaluation better captures the progress of the program. For example, selection intensity was not constant over the five cycles of selection. This highlights changes in decision processes, with selection of the best families on

Table 4. Evolution of the CNA6 population across cycles and the impact of selection on the three traits: grain yield (YLD, kg.ha^{-1}), days to flowering (FLW, days) and plant height (PHT, cm).

Trait	Cycle	Controls	Families		S	Mean Difference			RS	H^2
			Evaluated	Selected		Ev-Ctrl	Sel-Ctrl	Cycles		
YLD	3	3036.3	2317.7	2514.8 (47)	197.0	−718.5	−521.5	.	39.7	0.20
	4	2406.8	1823.7	1958.5 (45)	135.0	−583.0	−448.3	−556.3	20.0	0.15
	5	3890.9	3350.4	3434.6 (46)	84.2	−540.5	−456.3	1476.1	4.8	0.06
	6	3536.5	3301.6	3395.4 (25)	93.8	−234.9	−141.1	−39.2	9.7	0.10
	7	3038.1	4166.7	4417.3 (38)	251.0	1128.6	1379.2	1021.9	146.0	0.41
FLW	3	77	80.6	81.1 (47)	0.5	3.6	4.1	.	0.3	0.62
	4	87.3	92.8	92 (45)	−0.9	5.5	4.7	10.8	−0.6	0.73
	5	71.7	71.3	71.5 (46)	0.2	−0.4	−0.2	−20.5	0.1	0.56
	6	82.5	80.7	80.7 (25)	−0.08	−1.8	−1.8	9.2	−0.03	0.43
	7	79.2	82.5	82.2 (38)	−0.2	3.3	3.0	1.6	−0.2	0.47
PHT	3	102.1	102.1	102 (47)	0.4	0	−0.1	.	0.1	0.24
	4	90	96	95.9 (45)	−0.1	6	5.9	−6.5	−0.03	0.24
	5	113.5	116.1	116 (46)	−0.5	2.6	2.5	19.7	−0.07	0.15
	6	108.3	120.1	120 (25)	0.2	11.8	11.7	4.7	0.02	0.15
	7	104.4	105.5	106 (38)	0.4	1.1	1.4	−14.4	0.21	0.40

The following information is reported: mean for the controls, the S$_{0:2}$ families and selected families (the number of families selected is in brackets); selection differential between the selected families and all S$_{0:2}$ families for a given cycle (S); mean difference between evaluated families and controls (Ev-Ctrl); mean difference between the selected families and controls (Sel-Ctrl); mean difference between cycles (Cycles); and gain from selection (kg.ha^{-1}) with the selected families by cycle (RS); broad-sense heritability (H^2).

Table 5. Genotypic (above the diagonal) and phenotypic (below the diagonal) correlations between the traits grain yield (YLD, kg.ha^{-1}), plant height (PHT, cm) and days to flowering (DTF, days), by cycle and over the whole period.

Cycle	TRAIT	FLW	PHT	YLD
3	FLW	1	0.04	−0.05
	PHT	−0.07	1	0.10
	YLD	−0.19 **	0.55 **	1
4	FLW	1	0.13	−0.22 *
	PHT	−0.80 **	1	−0.10
	YLD	−0.89 **	0.86 **	1
5	FLW	1	−0.03	−0.13
	PHT	−0.50 **	1	0.31 **
	YLD	−0.07	0.21 **	1
6	FLW	1	−0.05	−0.26 *
	PHT	0.37 **	1	0.16
	YLD	0.01	0.62 **	1
7	FLW	1	0.04	0.12
	PHT	0.14 **	1	0.13
	YLD	0.21 **	0.80 **	1
All Cycles	FLW	1	−0.57 **	−0.20 **
	PHT	−0.40 **	1	0.30 **
	YLD	−0.22 **	0.51 **	1

different traits and the maintenance of genetic variability over the cycles. Indeed, the variation of selection intensity was associated with YLD performance of the families in a given cycle, but also with other traits of interest, such as plant architecture, lodging, earliness, disease resistance and grain quality, all of which are fundamental qualities for both rice growers and the agri-food industry. YLD evaluations were initiated in multisite trials in cycle

3, in which low-yielding families were easily detected and eliminated, raising the population mean. The gain observed for traits with a higher heritability (plant height and flowering time) in the S$_{0:2}$ families were also partly the results of gains achieved through selection in generations S$_0$ and S$_{0:1}$, as reported by Breseghello et al. (2009).

Drought tolerance is another objective of the upland rice breeding program. In cycle 7, the first selection criterion applied was a grain yield greater than 1500 kg.ha^{-1} under drought conditions, evaluated in managed stress trials, followed by grain yield in standard conditions. This choice of selection criteria was designed to exert a high selection pressure for drought tolerance in the CNA6 population, thereby improving the general response of the CNA6 population to stress in the next cycle.

Several other studies have analyzed the rate of genetic gain for YLD. The estimates obtained varied greatly, as a function of the breeding program concerned and its specific features (irrigated or upland rice, type of population, breeding objectives, time scales). In most studies, the rate of genetic gain for YLD was between 0.5% and 2% per year (Atroch and Nunes, 2000; Breseghello et al., 2011; Breseghello et al., 1999; da Silva et al., 2021; de Souza et al., 2007; Juma et al., 2021; Kumar et al., 2021; Moura et al., 2021; Muralidharan et al., 2002; Peng et al., 2000). In upland or rainfed rice breeding programs, the highest estimate was that reported by Soares *et al* (1999), who found a genetic gain of 46.4 kg.ha^{-1} (3.37%) for the intermediate-maturity group of the breeding program developed in Minas Gerais State using advanced lines. Conversely, the lowest rate was that reported by Moura et al. (2021) who found a negative genetic gain of −89.8 kg.ha^{-1} (−2.0%) per year over four years of a pedigree-based breeding program. According to the authors, this negative gain was due to an intensification of selection for grain quality traits. For irrigated rice, Peng et al. (2000) at the International Rice Research Institute (IRRI) used era trials with released varieties to estimate an annual gain of approximately 75 to 81 kg.ha^{-1}, equivalent to 1% for YLD. Lower values were obtained by Juma *et al* (2021) over a longer period (1964 to 2014): 8.8 kg.ha^{-1} to 17.4 kg.ha^{-1} per year.

Fig. 3 Distribution of phenotypic values for advanced lines from the product development pipeline. The data is originated from the EP (preliminary multisite trials), ER (regional multisite trials) and VCU (multisite trials of value for cultivation and use). The lines derived from the CNA6 population are shown in orange, other lines are in blue, and controls are in black. The vertical dashed lines indicate the mean value.

Genetic gains for yield have been low since the Green Revolution, as the IR8 cultivar set a high reference yield, with the introduction of a semi-dwarf architecture. Studies estimating long-term gains tend to report a lower relative gain (in %). Muralidharan et al. (2002), for example, quantified the genetic gain in India's national rice breeding program from 1995 to 2013 and found that there had been no genetic gain. Similarly, Santos et al. (1999) analyzed data for 22 years of the breeding program in Minas Gerais State and found non-significant gain. One exception is the study by Breseghello et al. (2011), who evaluated the genetic gain resulting from Embrapa's upland rice breeding program over a period of 25 years, from 1984 to 2009, and found an annual gain was 45.0 kg.ha^{-1} (1.44%). Given the diversity of the breeding program and of estimates for the genetic gain in grain yield, it is difficult to draw clear conclusions about the determinants of successful breeding strategies. However, shorter breeding cycles, such as those currently applied to the CNA6 population, can efficiently deliver high rates of genetic gain over long time period.

Plant height and flowering time. YLD and PHT were considered more important than FLW in the first few cycles of selection applied to the CNA6 population. FLW has become more important in recent years, with the perspective of inserting upland rice into sustainable integrated production systems, in which a short cycle is desirable to ensure that different crops are grown in the same

area in a single year. The upland rice breeding program aims to develop cultivars with a PHT between 90 and 100 cm, considering that taller plants tend to lodge, especially if fertility and sowing density are high. Other studies investigating the genetic trend for PHT have generally reported stability or a significant decrease (Breseghello et al. 2011; Breseghello et al. 2009; de Souza et al. 2007; Morais Júnior et al. 2017).

The objective of the Embrapa upland rice breeding program is to develop cultivars with a time to flowering of 50–60 days. However, instead of maintaining or reducing the FLW, we found that five cycles of selection increased FLW, by 1.5 days per cycle. However most of this trend was related to non-genetic factors. Breseghello et al. (2009) also reported an annual increase of 0.6 days (0.8%) for the CG3 population. By contrast, Morais Júnior et al. (2017) observed a genetic gain of −6.4 days per cycle (−1.3% per year) for the CNA12 population. The correlations obtained in this study highlight the complexity of multi-trait selection. The use of a well-defined selection index might make it possible to avoid conflicts between the genetic gains for different traits in future cycles (Baker, 2020; Céron-Rojas and Crossa, 2018).

Importance of breeding decisions
Changes in the crossing strategy. By contrast to previous cycles, recombination between selected families was achieved by manual

crossing in cycle 6, to generate the population for cycle 7. The decision to change crossing strategy, from the genetic male sterility approach to manual crosses, was due to an observed increase in PHT in the population, despite negative selection for this trait. Pollen dispersal from taller plants is more effective than from shorter plants, favoring taller plants over shorter ones (Morais Júnior et al., 2017). Thus, despite selection for shorter plants, during recombination steps, taller plants are more likely to pass their alleles on to offspring for the next cycle of RS. Male sterility was therefore abandoned for the generation of the population for cycle 7. However, in a program where the number of manual crosses is a limiting factor, manual removal of the tallest plants from the recombination field each cycle may be an alternative.

This change in methodology had a significant impact on the genetic variability of the CNA6 population. The genetic variance of families for YLD in cycle 7 was almost five times that in cycle 6 (Table S4). The manual crossing of the selected families from each subpopulation in a circular diallel scheme, associated with the use of an appropriate number of parents, improves control over population inbreeding, thus helping to maintain genetic diversity. Manual crosses were planned according to the divergence and complementarity between parents. In a study in which SSR markers were used in the first two cycles of RS, Pinheiro et al. (2012) observed a maintenance of genetic diversity in the CNA12 population when a manual crossing strategy was used, contrasting with the loss of diversity observed when random crossing based on male sterility was used.

Manual crossing also makes it possible for breeders to trace the pedigree within the population, which can be used in mixed models to increase the prediction accuracy of breeding values (Piepho et al., 2008). A knowledge of pedigrees can also enable breeders to make better decisions throughout the process, from the planning of crosses to the release of new cultivars for a target environment. Thus, connecting the genotypes in historical datasets through pedigrees renders phenotypic data more useful (Breseghello et al., 2021). This is crucial for single-step genomic prediction models and can facilitate the implementation of genomic selection in breeding programs (Legarra et al., 2014; Morais Júnior et al., 2018b).

Change in trial design for progeny testing. The first four cycles of selection were performed with augmented block design, with no within-site repetition. In the seventh cycle, an alpha-lattice with two repetitions was used. The augmented block design resulted in a lower precision for the phenotypic evaluation, consequently decreasing the gain due to selection. The gain in precision achieved with the lattice design resulted in a two to six-fold increase in heritability for YLD. Augmented block designs are widely used for early-stage evaluations in breeding programs, for which the number of genotypes is high (Clarke and Stefanova, 2011; Zystro et al., 2018). At early stages, breeders usually favor the number of genotypes over the number of replicates for a given number of plots. However, the allocation of resources between genotypes, controls and replicates is always a tricky issue to solve, because more controls or replicates generally result in a better estimation of phenotypic performance (Mead, 1997). In this study, the observed benefits of the lattice design with two replicates support the change in experimental design to favor a greater genetic gain. The lattice design is now the standard design for multisite trials with $S_{0:2}$ families in the Embrapa rice breeding program.

Connectivity between sites and years is a key aspect in the estimation of realized genetic gain in breeding programs (Rutkoski, 2019). The maintenance of benchmark varieties for the $S_{0:2}$ trials is therefore essential, to ensure that the analysis is robust, as in this study. The use of varieties as controls in multisite trials provides a link between different cycles, facilitating the evaluation of genetic gain. However, varieties that are grown for many years tend to become more susceptible to disease, which may prevent the meaningful phenotyping of other important traits and increase the magnitude of the nongenetic trend (Fischer and Edmeades, 2010; Mackay et al., 2011; Piepho et al., 2014). It was for this reason that BRS Esmeralda, released in 2013, was introduced as a control in cycle 7, whereas BRS Primavera, which was present in the last five cycles of selection will not be included as a control in cycle 8. The comparison of population means with relevant controls over several years is essential. For example, in cycle 7, the mean YLD of the CNA6 population exceeded that of most of the controls, including BRS Esmeralda, a high-yielding variety released in 2013 (Colombari Filho et al., 2013). The selected families were 45% more productive than the controls, revealing a considerable potential for the eighth cycle of breeding.

Integration of recurrent selection and product development. The potential of the CNA6 population to provide high-quality variation for the product development pipeline is demonstrated by the presence of lines originally from CNA6 in the VCU trial, the last step in the upland rice breeding program. These elite lines were generated from CNA6 families used as parents in crosses combining traits to generate segregating populations and, subsequently, lines with a high potential. These lines were tested in multisite trials of the upland rice breeding network in Brazil for several generations, with an increasing level of selection pressure. At the last stage, lines are selected not only for YLD, PHT and FLW, but also for grain quality, disease resistance, lodging and water deficiency tolerance. The presence of these lines in the VCU trials, after several generations of increasing selection pressure demonstrates the potential of these lines to become new cultivars for release into the market. The increase in the potential of the population to generate superior lines observed here results from a displacement of the population mean or genetic gain in the desired direction. For example, the families selected in cycle 7 were 45.4% more productive than the controls, including BRS Esmeralda, the last cultivar released from the Embrapa upland rice breeding program. Fifteen families from cycle 7 of the CNA6 population were used as new parents in crosses for the product development pipeline.

Next steps in the breeding strategy

In RS programs, the rate of genetic gain can be increased by accelerating the breeding cycles, intensifying the selection pressure, improving the precision of the evaluation, or by any combination of such changes (Bernardo, 2020; Cobb et al., 2019). Off-season nurseries can be used to reduce the cycle time in plant breeding. Nonetheless, in RS schemes based on progeny testing, it is advisable to grow more than one crop per year only for the recombination step and seed multiplication, because plant and progeny selection must be performed during the normal crop season to ensure that the conditions are representative of the target environments (Heinemann et al., 2015). The constraints on the use of off-season nurseries and selection pressure intensification imposed by the evaluation of a larger number of families can be overcome by the use of modern technologies. Population improvement by RS is highly flexible and provides opportunities for the combination of classical genetic improvement tools with molecular genetics or genomics (Bartholomé et al., 2022; Martínez et al., 2014). This integration of innovative technologies into the process has a considerable potential to render rice breeding faster and more effective, with a more intensive exploration of the tremendous genetic diversity present within this species (Wang et al., 2018). One potential alternative for the acceleration of genetic gains would be the use of genomic recurrent selection (GRS), as recently evaluated (Baertschi et al., 2021; Morais Júnior et al., 2018a; Spindel et al., 2015). This approach can accelerate the gain due to selection per unit time by greatly shortening the

breeding cycle and through the accurate selection of plants at early stages (e.g., S_0 or $S_{0:1}$ generations) for immediate recombination (Gaynor et al., 2017). Recurrent genomic selection would also make it possible to apply a high selection intensity, given that large populations can be developed and genotyped in a short time and at reasonable cost. This approach therefore has the potential to promote efficient and cost-effective breeding programs for self-pollinated species, such as rice, as demonstrated in other studies (Ahmadi et al., 2020; Bartholomé et al., 2022). The selection of families with a high potential by GRS will also benefit the product development pipeline (Gaynor et al., 2017). The best families would be identified even earlier, making it possible to use them directly for product development, thereby increasing the chances of developing high-value cultivars for commercial use. The integration of GRS in the Embrapa rice breeding program should make it possible to increase the rate of genetic gain in the future.

CONCLUSION AND PROSPECTS

The application of five cycles of RS to the CNA6 population resulted in significant genetic gains for grain yield. The integration of this strategy into a two-phase breeding program also led to improvement of quantitative traits while selecting for other traits of interest, such as grain quality, disease resistance and pest resistance.

The efficiency of the RS method can be further improved by the integration of molecular markers to increase selection intensity and to reduce cycle length, to ensure the faster delivery of more diverse material with better performances to the product development pipeline.

DATA AVAILABILITY

The data used in this study are available at the following address: https://doi.org/10.5061/dryad.1g1jwsv28.

REFERENCES

Ahmadi N, Bartholomé J, Tuong-Vi C, Grenier C (2020). Genomic selection in rice: empirical results and implications for breeding. CABI: Wallingford, pp 243–258

Atroch AL, Nunes GHD (2000) Genetic progress in lowland rice in the state of Amapa, Brazil. Pesqui Agropecu Bras 35(4):767–771

Baertschi C, Cao T-V, Bartholomé J, Ospina Y, Quintero C, Frouin J et al. (2021) Impact of early genomic prediction for recurrent selection in an upland rice synthetic population. *G3 Genes|Genomes|Genetics* 11(12):1–16

Baker RJ (2020). *Selection indices in plant breeding*. CRC Press

Barros MS, Morais Júnior OP, Melo PGS, Morais OP, Castro AP, Breseghello F (2018) Effectiveness of early-generation testing applied to upland rice breeding. Euphytica 214(4):61

Bartholomé J, Prakash PT, Cobb JN (2022). Genomic prediction: progress and perspectives for rice improvement. In: Ahmadi N and Bartholomé J (eds) *Complex Trait Prediction: Methods and Protocols*. Springer US: New York, NY, pp 569–617

Bernardo R (2003) On the effectiveness of early generation selection in self-pollinated crops. Crop Sci 43(4):1558–1560

Bernardo R (2020) *Breeding for quantitative traits in plants*. 3. ed. Woodbury:Stemma Press

Breseghello F, de Mello RN, Pinheiro PV, Soares DM, Lopes Júnior S, Nakano, Rangel PH et al. (2021) Building the Embrapa rice breeding dataset for efficient data reuse. Crop Sci 61(5):3445–3457

Breseghello F, de Morais OP, Pinheiro PV, Silva ACS, da Maia de Castro E, Guimarães ÉP et al. (2011) Results of 25 years of upland rice breeding in Brazil. Crop Sci 51(3):914–923

Breseghello F, Morais O, Castro E, Prabhu A, Bassinello P, Pereira J et al. (2009) Recurrent selection resulted in rapid genetic gain for upland rice in Brazil. Int Rice Res News 34(4):1-4

Breseghello F, Rangel PHN, De Morais OP (1999) Yield gain through irrigated rice breeding in the Northeast Brazil. Pesqui Agropecu Bras 34(3):399–407

Ceccarelli S (2015) Efficiency of plant breeding. Crop Sci 55(1):87–97

Céron-Rojas JJ, Crossa J (2018). *Linear selection indices in modern plant breeding*. Springer Nature

Clarke GPY, Stefanova KT (2011) Optimal design for early-generation plant-breeding trials with unreplicated or partially replicated test lines. Aust NZ J Stat 53(4):461–480

Cobb JN, Juma RU, Biswas PS, Arbelaez JD, Rutkoski J, Atlin G et al. (2019) Enhancing the rate of genetic gain in public-sector plant breeding programs: lessons from the breeder's equation. Theor Appl Genet 132(3):627–645

Colombari Filho JM, de Resende MDV, de Morais OP, de Castro AP, Guimaraes EP, Pereira JA et al. (2013) Upland rice breeding in Brazil: a simultaneous genotypic evaluation of stability, adaptability and grain yield. Euphytica 192(1):117–129

da Silva AC, Carneiro VQ, dos Santos IG, Rosado RDS, Cruz CD, Soares PC (2021) Genetic progress over twenty-three years of irrigated rice breeding in southeastern Brazil. Acta Sci Agron 43(1):1–12

de Souza MA, de Morais OP, Heran REC, Cargnin A, Pimentel AJB (2007) Genetic progress of upland rice between 1950 and 2001. Pesqui Agropecu Bras 42(3):371–376

Dudley JW (2007) From Means to QTL: The Illinois long-term selection experiment as a case study in quantitative genetics. Crop Sci 47(S3):S-20–S-31

Eberhart SA (1964). Least Squares method for comparing progress among recurrent selection methods1. *Crop Sci* 4(2): cropsci1964.0011183X000400020036x

Federer WT, Raghavarao D (1975) On augmented designs. Biometrics 31(1):29–35

Fischer RA, Edmeades GO (2010) Breeding and cereal yield progress. Crop Sci 50(S1):S-85–S-98

Fujimaki H (1979) Recurrent selection by using genetic male sterility for rice improvement. JARQ 13(3):153–156

Gaynor RC, Gorjanc G, Bentley AR, Ober ES, Howell P, Jackson R et al. (2017) A two-part strategy for using genomic selection to develop inbred lines. Crop Sci 57(5):2372–2386

Grenier C, Cao TV, Ospina Y, Quintero C, Chatel MH, Tohme J et al. (2015) Accuracy of genomic selection in a rice synthetic population developed for recurrent selection breeding. Plos One 10(8):1–25

Hallauer AR, Carena MJ, Miranda Filho JD (2010). *Quantitative genetics in maize breeding*, Vol 6. Springer Science & Business Media

Hallauer AR, Darrah LL (1985) Compendium of recurrent selection methods and their application. Crit Rev Plant Sci 3(1):1–33

Heinemann AB, Barrios-Perez C, Ramirez-Villegas J, Arango-Londoño D, Bonilla-Findji O, Medeiros JC et al. (2015) Variation and impact of drought-stress patterns across upland rice target population of environments in Brazil. J Exp Bot 66(12):3625–3638

Juma RU, Bartholomé J, Thathapalli Prakash P, Hussain W, Platten JD, Lopena V et al. (2021) Identification of an elite core panel as a key breeding resource to accelerate the rate of genetic improvement for irrigated rice. Rice 14(1):92

Kumar A, Raman A, Yadav S, Verulkar SB, Mandal NP, Singh ON et al. (2021) Genetic gain for rice yield in rainfed environments in India. Field Crops Res 260:107977

Laidig F, Piepho H-P, Drobek T, Meyer U (2014) Genetic and non-genetic long-term trends of 12 different crops in German official variety performance trials and on-farm yield trends. Theor Appl Genet 127(12):2599–2617

Legarra A, Christensen OF, Aguilar I, Misztal I (2014) Single Step, a general approach for genomic selection. Livest Sci 166:54–65

Mackay I, Horwell A, Garner J, White J, McKee J, Philpott H (2011) Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time. Theor Appl Genet 122(1):225–238

Martínez CP, Torres EA, Chatel M, Mosquera G, Duitama J, Ishitani M et al(2014) 5 rice breeding in Latin America Plant Breed Rev 38:187–278

Mead R (1997). Design of plant breeding trials. In: Kempton RA, Fox PN and Cerezo M (eds) *Statistical Methods for Plant Variety Evaluation*. Springer Netherlands: Dordrecht, pp 40–67.

Morais Júnior OP, Breseghello F, Duarte JB, Coelho ASG, Borba TCO, Aguiar JT et al. (2018a) Assessing prediction models for different traits in a rice population derived from a recurrent selection program. Crop Sci 58(6):2347–2359

Morais Júnior OP, Breseghello F, Duarte JB, Morais OP, Rangel PHN, Coelho ASG (2017) Effectiveness of recurrent selection in irrigated rice breeding. Crop Sci 57(6):3043–3058

Morais Júnior OP, Duarte JB, Breseghello F, Coelho ASG, Morais OP, Magalhães Júnior AM (2018b) Single-step reaction norm models for genomic prediction in multienvironment recurrent selection trials. Crop Sci 58(2):592–607

Morais Júnior OPD, Melo PGS, Morais OPD, Castro APD, Breseghello F, Utumi MM et al. (2015) Genetic progress after cycles of upland rice recurrent selection. Sci Agric 72:297–305

Morais OPD (2005). Genetic mean and variability in recurrent selection. In: Guimaraes EP (ed) *Population improvement: A way of exploiting the rice genetic resources of Latin America*. Food and Agriculture Organization of the United Nations (FAO): Rome

Moura AMD, Botelho FBS, Tomé LM, Rodrigues CS, Silva CSCD, Silva MPD (2021) Genetic progress in upland rice breeding program for grain yield and plant height. *Ciência e Agrotecnologia* 45(1):1–8

Muralidharan K, Prasad GSV, Roa CS (2002) Yield performance of rice genotypes in international multi-environment trials during 1976–97. Curr Sci 83(5):610–619

Peng S, Laza RC, Visperas RM, Sanico AL, Cassman KG, Khush GS (2000) Grain yield of rice cultivars and lines developed in the Philippines since 1966. Crop Sci 40(2):307–314

Piepho H-P, Laidig F, Drobek T, Meyer U (2014) Dissecting genetic and non-genetic sources of long-term yield trend in German official variety trials. Theor Appl Genet 127(5):1009–1018

Piepho HP, Möhring J, Melchinger AE, Büchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161(1):209–228

Pinheiro LDS, Rangel PHN, Vianello RP, Brondani C (2012) Genetic variability of rice recurrent selection populations as affected by male sterility or manual recombination. Pesqui Agropecu Bras 47::808–814

Ramirez-Villegas J, Heinemann AB, Pereira de Castro A, Breseghello F, Navarro-Racines C, Li T et al. (2018) Breeding implications of drought stress under future climate for upland rice in Brazil. Glob Change Biol 24(5):2035–2050

Rangel PHN, de Morais OP, Zimmermann FJP (2002) Grain yield gains in three recurrent selection cycles in the CNA-IRAT 4 irrigated rice population. Crop Breeding Applied Biotechnology 2(3):369–374

Rutkoski JE (2019). A practical guide to genetic gain. In: Sparks DL (ed) *Advances in Agronomy*. Academic Press. Vol. 157, pp 217–249

Santos PG, Soares PC, Soares AA, Morais OPD, Cornélio VMDO (1999) Evaluation of the genetical progress obtained in 22 years on the improvement of irrigated rice in Minas Gerais, Brazil. Pesqui Agropec Bras 34::1889–1896

Singh R, Ikehashi H (1981) Monogenic male-sterility in rice: induction, identification and inheritance 1. Crop Sci 21(2):286–289

Soares AA, Santos PG, De Morais OP, Soares PC, Reis MD, De Souza MA (1999) Genetic progress obtained by upland rice breeding in twenty one years of research in the state of Minas Gerais, Brazil. Pesqui Agropecu Bras 34(3):415–424

Spindel J, Begum H, Akdemir D, Virk P, Collard B, Redoña E et al. (2015) Genomic selection and association mapping in rice (oryza sativa): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLoS Genet 11(2):e1004982

Taillebois J, Guimaraes E. (1989) *International Symposium on Hybrid Rice, Changsha, Hunan (China), 6–10 Oct 1986*. IRRI

UNITED NATIONS (2017) Department of Economic and Social Affairs. World population projected to reach 9.8 billion in 2050, and 11.2 billion in 2100

van Ginkel M, Ortiz R (2018) Cross the best with the best, and select the best: HELP in breeding selfing crops. Crop Sci 58(1):17–30

Walsh B, Lynch M (2009). Evolution and selection of quantitative traits. *Foundations Sunderland, MA: Sinauer*

Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z et al. (2018) Genomic variation in 3010 diverse accessions of Asian cultivated rice. Nature 557(7703):43–49

Zystro J, Colley M, Dawson J (2018). Alternative experimental designs for plant breeding *Plant Breed Rev*, 87–117

AUTHOR CONTRIBUTIONS

The experiments were performed by AC, FB, IVF, MMU and JAP. TVC, AC and JB performed the analysis. AC and JB wrote the first draft of the manuscript, which was further edited by all the co-authors. AC and JB designed and coordinated the study. All the authors have read and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41437-023-00636-3.

Correspondence and requests for materials should be addressed to Adriano Pereira de Castro.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.