FUNCTIONAL PLANT BREEDING JOURNAL
Scientific Journal. ISSN 2595-9433
Volume 6, Article n. 05, Jan/December
D.O.I. http://dx.doi.org/10.35418/2526-4117/v6a05
Received: 12/12/2023 - Accepted: 03/02/2024

Comparison of grain yield and identification of QTLs of rice populations derived from the Bulk and SSD (Single Seed Descent) methods

Mariana Rodrigues Feitosa Ramos[1,2], João Antônio Mendonça[1], Rosana Pereira Vianello[1], Claudio Brondani[1,*]

[1] *Embrapa Arroz e Feijão, Santo Antônio de Goiás, Goiás, Brazil;*
[2] *Universidade Federal de Goiás, Escola de Agronomia, Goiânia, Goiás, Brazil.*
** Corresponding author: claudio.brondani@embrapa.br*

Abstract: Bulk and SSD are widely used methods for developing segregating populations in plant genetic improvement programs. This work aimed to analyze grain yield (GY) QTLs using two populations from the Epagri 108 × Irat 122 cross, advanced by the SSD (F_8 generation) and Bulk ($F_{7:8}$ generation) methods. The 158 recombinant inbred lines (RILs) of each method (SSD and Bulk) were evaluated for two years (2017 and 2018 crop seasons), in an 18x18 lattice square experimental design with two replications. The RILs were genotyped using the DArTseq® methodology, which generated around 6 thousand SNPs and SilicoDArTs. The statistical model adopted for the analysis of GY data was a mixed linear model (MLM). For the first and second year evaluations and joint analysis, the RILs-Bulk group presented higher productivity averages when compared to the group RILs-SSD and Checks. However, in relation to the genetic variance component, the SSD group presented the highest estimate followed by Bulk and Checks. The GY of the RILs-Bulk ranged from 4,010.75 kg.ha^{-1} to 5,815.42 kg.ha^{-1}, while the RILs-SSD ranged from 3,321.76 kg.ha^{-1} to 8,096.27 kg.ha^{-1}, both exceeding the Check group, which ranged from 2,754.30 kg.ha^{-1} to 3,643.73 kg.ha^{-1}. For QTL analysis, multiple interval mapping was applied, with 2,115 SNPs genotyped in the RILs-SSD, resulting in three QTLs identified (qGYLD6, qGYLD6.2 and qGYLD9). For RILs-Bulk, with 2,354 markers, three QTLs were identified (qGYLD6 and qGYLD9), and in the joint analysis, one QTL was mapped (qGYLD7). The next step will be to validate the markers identified by genotyping accessions from the Germplasm bank before being used in assisted selection for the GY trait. Both the SSD and Bulk methods were efficient in obtaining high grain yield RILs and in the ability to identify QTLs. However, the Bulk method was more practical to handle and cheaper to obtain RILs compared to SSD.

Keywords: Recombinant Inbred Lines, SNP markers, Interval Mapping, grain yield.

Introduction

Different breeding methods for creating segregating populations can be used in rice, an autogamous species, with the aim of obtaining recombinant inbred lines superior to those already in cultivation, such as the single seed descent (SSD) and the population (Bulk) methods (Ramalho et al., 2012). The Bulk method consists of collecting all plant seeds in each generation and planting a random sample of seeds to propagate the next generation (Haddad and Muehlbauer, 1981). Due to inequality in fertility, highly competitive

variants are represented at higher frequencies in successive generations, resulting in decreased variability in the population and increased frequencies of closely related genotypes. The Bulk method allows the conduction of a greater number of segregating populations of interest for breeding, with higher grain yield averages and lower costs involved in the evaluation of field experiments (Mendes et al., 2011). The SSD method consists of removing a single seed from each plant, starting from F2, and sowing the seeds to propagate in the next generation (Silva et al., 2013). As a result, the SSD method allows obtaining homozygous lines with greater control of the segregating population; however, it can reduce the number of crosses evaluated, as it requires more space for the cultivation of each progeny.

Recombinant inbred lines, with a high degree of homozygosity after successive generations of self-fertilization, can also be used for molecular genetic studies. The genotyping of RILs allows the determination of the average genetic distance between RILs, and the identification of genes, or gene blocks, associated with characters of interest via quantitative trait loci (QTL) analysis. QTL loci are chromosomal segments that govern quantitative traits, but this definition refers only to a statistical association between a region of the genome and a phenotypic character, which have a continuous distribution, such as plant height, grain yield, etc. (Guo et al. 2014). For characters with low heritability, phenotypic selection is carried out in more advanced generations, as the heritability and statistical precision of estimating progeny averages increase with the increase in the number of repetitions, generations, locations and years of testing, which leads to a very large increase in the number of plants to be evaluated (Dixit et al., 2019).

The grain yield is regulated by the interaction of several genes with small effects and by environmental action on the genotype, and tends to have low or moderate heritability (Xing and Zhang, 2010). QTL mapping and cloning studies have been important in identifying genes that regulate grain weight, including both genes for grain size and grain filling, for example (Qi et al., 2012; Dixit et al., 2019). In this study, RILs from the cross Epagri 108 (*Oryza sativa* ssp. Indica) × Irat 122 (*Oryza sativa* ssp. Japonica), advanced through the SSD and Bulk methods, were genotyped by SNPs and SilicoDArTs markers and evaluated in field experiments. The objectives of this study were to compare the RILs developed by the two methods in terms of genetic variability and productive performance, and the ability of these populations to identify QTLs related to grain yield.

Material and Methods

Advancement of Bulk and SSD Populations

The cross Epagri 108 (*O. sativa* ssp. Indica) × Irat 122 (*O. sativa* ssp. Japonica) was carried out in the year 2005. The rice cultivar for irrigated cultivation system, Epagri 108, was commercially released in 1996 and exhibits broad adaptability for cultivation throughout Brazil. The cultivar for upland cultivation system, Irat 122, was released in 1980 by CIAT, introduced in Brazil, and maintained in the Embrapa's Germplasm bank. The Epagri 108 x Irat 122 population was advanced through the Bulk method starting from the F_2 generation in 2006. For the generation advancement, samples of 500 seeds were taken and planted again the following year, until the F_7 generation, where plants were randomly selected, and each giving rise to an $F_{7:8}$ family. These families were arranged in five-meter rows, and seeds were harvested from each family. The Epagri 108 × Irat 122 population was also advanced through the SSD (single seed descent) method from 2006. Starting from a population of 500 seeds in F_2, ten seeds from each F_2 plant were sown, and after 20 days, five seedlings were transplanted into the field. From these five plants, only one was randomly chosen to advance to the next generation. This entire process was repeated until the F_7 generation, when seeds (F_8) were harvested from all plants of each family. For the two years of experiments (2017 and 2018) at Palmital

farm in Goianira-GO, 158 Bulk lines ($F_{7:8}$) and 158 SSD lines (F_8) were randomly selected to be evaluated in the field experiments.

Assessment of advanced families by Bulk and SSD

The SSD and Bulk populations resulting from the cross Epagri 108 × Irat 122 were evaluated in two field experiments conducted during the 2016/2017 and 2017/2018 growing seasons at Palmital Farm, located in Goianira municipality, Goiás, Brazil (16°30'23''S; 49°17'00''W; 823 m). In each experiment, 158 RILs from each method, along with 8 checks (two parents and the cultivars Araguaia, Maninjau, Irga 417, BRS Catiana, BRS Tropical, INTA Puitá CL), were evaluated using an 18x18 lattice square experimental design with plots consisting of 4 rows of 3 meters, with two repetitions. The evaluated trait was grain yield (GY), with all plants from the two central rows of each plot being harvested and weighed using a precision scale. The weight values in grams were then converted to $kg.ha^{-1}$.

Statistical Analysis Using Mixed Models

The adopted statistical model for data analysis was a mixed-effects model:

(Eq.1)

$$Y_{ijkmn} = \mu + g_{i(m)} + r_j + b_{k(j)} + p_m + t_n + \varepsilon_{ijkmn}$$

Where Y_{ijkmn} is the observed value of genotype i, in repetition j, within block k, and μ is the constant inherent to all observations, representing the overall mean of the trial. As random effects in the model, we have: $g_{i/m}$ is the effect of genotype i within group m, r_j is the effect of repetition j, $b_{k/j}$ is the effect of block k within repetition j, and ε_{ijkmn} is the effect of associated errors. Similarly, as fixed effects of the model, we have: p_m which is the effect of groups containing genotype i, and t_n, which is the effect of checks n.

The effects of the model were estimated using the Smmer package in the R software (R Core Team, 2022). The variance components were estimated using the restricted maximum likelihood (REML) method. This allowed obtaining the best unbiased estimate for fixed effects, i.e., the empirical best linear unbiased estimators (eBLUEs), and the best unbiased linear prediction for random effects, the empirical best linear unbiased predictors (eBLUPs). The significance of each effect within the model was tested both in individual analyses and in the joint analysis using the analysis of deviance. The analysis was performed by disregarding each of the effects under test (repetition, block, genotype) and subtracting the deviance from the model without the respective effect from the deviance of the complete model. The significance of each effect was verified using the likelihood ratio test (LRT), using the χ^2 distribution at 1% and 5% probability levels, with one degree of freedom.

The coefficient of experimental variation (CVe) was determined by the formula:

(Eq.2)

$$\widehat{CV}_e = \frac{\hat{\sigma}_e^2}{m} \text{ x } 100$$

The selective accuracy was determined by the formula:

(Eq.3)

$$\hat{r}_{gg} = [1 - \frac{1}{1 + bCV_r^2}]^{1/2}$$

Molecular analysis by SNP markers

Genomic DNA was extracted from young leaves of the inbred lines and their parents using the commercial kit DNeasy 96 Plant Kit (Qiagen). After sample preparation, DNA was sent to Diversity Arrays Technology (DArT) Pty Ltd (Bruce, Australia) for genotyping. The DArTseq methodology generated dominant SilicoDArTs markers (presence/absence) and SNPs (Single Nucleotide Polymorphism). The average genetic distance between the RILs of each Bulk and SSD methods was obtained by the GenAlEx version 6.51b2 program (Peakall & Smouse, 2012) using the distance coefficient of Smouse & Peakall (1999).

LD extension estimates were calculated by the usual method (r^2) using the LDcorSV package in R (Desrousseaux et al., 2013). In contrast, the Genetic Relationship Matrix (GRM) was estimated by the algorithm proposed by Yang et al., (2010), using the GCTA package (Yang et al., 2011). The LD decay was explained by the nonlinear model (Hill & Weir, 1988) and adjusted to the "nls" function in the R program. The haplotype blocks were identified using the Haploview software (Barret et al., 2005), with a minimum percentage of genotyping of individuals ≥ 75% and $R^2 \geq 0.8$.

The linkage map of each RIL population (SSD and Bulk) was obtained using the MapDisto program version 1.7.7 (Lorieux, 2012), using LOD ≥ 3 and maximum recombination frequency $\theta = 0.3$. The recombination fractions were transformed into genetic distances (cM) by the Kosambi function. QTL analysis was performed using the statistical software R and the R/qtl package, using SNP markers and adjusted means of RILs (Blups) for the traits grain yield and plant height. The multiple QTL mapping strategy was used, with a LOD value considered equal to or greater than three. The nomenclature of QTLs followed the guidelines described by McCouch et al. (1997).

Results

It was verified by the Deviance analysis that for the grain yield (GY) character the effects of the genotypes from the first year of experimentation (2017), only the SSD method was significant ($p<0.05$) and, consequently, its genetic variance component, while the Bulk and the checks showed no statistical significance. For the second year (2018), the three groups of genotypes (Bulk method, SSD method and checks) were significant ($p<0.01$), as well as their genetic variance components. For the joint analysis of the two years, only SSD was significant ($p<0.01$). Deviance analysis thus demonstrated the presence of genetic variability among the strains evaluated (Table 1).

Table 1. Deviance analysis data for the RILs-Bulk, RILs-SSD and checks genotype groups for the grain yield (GY) trait.

	GY - 2017			GY - 2018			GY – Joint Analysis		
	LRT	p.value	σ^2g	LRT	p.value	σ^2g	LRT	p.value	σ^2g
RILs-Bulk	2.52	0.11^{ns}	901724.8	39.68	2.99e-10*	281726.1	2.38	0.12^{ns}	221473.2
RILs-SSD	3.92	0.05*	1144089	116.85	0*	697206.8	9.17	0.002*	459966.8
Checks	0.069	0.79^{ns}	603289.2	16.79	4.16e-05*	911485	1.25	0.26^{ns}	806471.6

σ^2g = genetic variance; * significant at 5% probability and ^{ns} not significant by the LRT test.

With the average values adjusted by BLUEs (fixed effects) for the grain yield character, the Bulk group stood out with the highest averages (4,953.25 kg.ha^{-1}; 3,225.59 kg.ha^{-1} and 4,189.42 kg .ha^{-1}), respectively for the 1st and 2nd year and joint analysis. The SSD group had the second best productive performance, with averages (4,284.43 kg.ha^{-1}; 2,735.15 kg.ha^{-1} and 3,614.56 kg.ha^{-1}) for the 1st and 2nd year and joint analysis. The checks had lower averages than the Bulk and SSD groups (3,158.01 kg.ha^{-1}; 3,369.85 kg.ha^{-1} and 3,456.70 kg.ha^{-1}) for the 1st and 2nd year and joint analysis (Table 2). Thus, the RILs of both methods were superior to the check cultivars, indicating that, for the locations and years evaluated, they are candidates for the development of future commercial rice cultivars.

Table 2. Adjusted means (Blue - fixed effect) of the RILs-Bulk, RILs-SSD and Checks groups for grain yield (GY).

	Grain Yield (GY)		
	2017	2018	Joint Analysis
RILs- Bulk	4,953.25*	3,225.59*	4,189.42*
RILs-SSD	4,284.34*	2,735.15*	3,614.56*
Checks	3,158.01*	3,369.85*	3,456.70*

* Significant at 5% probability by the F test.

The large magnitude of genetic variance (σ^2g) indicates that there is genetic variability within the population of RILs eval-

uated in the Bulk (RILs-Bulk) and SSD (RILs-SSD) methods and in the checks. In the first year of the experiment for the GY trait, only the SSD group showed a significant σ^2g component ($p<0.05$) in the LRT analysis, indicating greater genetic variability within the group of 158 RILs-SSD. In the second year of the experiment, for the GY trait, the RILs of both methods and checks were significant by the LRT test, but the checks showed higher σ^2g, followed by RILs-SSD and finally, RILs-Bulk. For the joint analysis, only the SSD group was significant. Consistently, a higher h^2 was also observed in SSD (0.23; 0.77; 0.29) than in Bulk (0.19; 0.57; 0.17) respectively for the 1st year and 2nd year and joint analysis (Table 3).

Experimental precision has a large effect on estimating the general genetic parameters of a population. The experimental coefficients of variation (CVe) of this work indicate that the two years of experiment had good to moderate experimental precision for the trait evaluated (Table 3). The CVe in the 1st year of experiment was high (59.46), and in the 2nd year it was within expectations (21.30). The accuracy of the present work was moderate to high for the trait evaluated, with the experiments considered to have good precision (Table 3).

Table 3. Estimates of genetic parameters genetic variance (σ^2g), heritability (h^2), selective accuracy (r_{gg}); environmental variance (σ^2e) and environmental variation coefficient (CVe) of RILs-Bulk, RILs-SSD and Checks for the trait Grain Yield (GY).

	Grain Yield (GY)								
	2017			2018			Joint Analysis		
	RILs-Bulk	RILs-SSD	Checks	RILs-Bulk	RILs-SSD	Checks	RILs-Bulk	RILs-SSD	Checks
σ^2g	901,724.8	1,144,089.0	603,289.2	2,817,261.0	6,972,068.0	911,485.0	2,214,732.0	4,599,668.0	806,471.6
h^2	0.19	0.23	0.14	0.57	0,77	0.81	0.17	0.29	0.42
r_{gg}		0.48			0.85			0.49	
σ^2e		739,445.7			410,779.0			431,570.0	
CVe (%)		59.46			21.30			52.25	

Genetic Mapping

Genotyping of RILs-Bulk and RILs-SSD and the two genitors using the DArTseq platform resulted in the identification of SNPs (codominant) and SilicoDArTs (dominant) markers. Genetic mapping distributed the markers across the 12 rice chromosomes, and the number of polymorphic markers varied between RILs-Bulk and RILs-SSD populations, because of the filtering performed (Table 4). The no distorted segregation markers identified by the $_X^2$ test which fit the expected F_8 segregation of 1:1, resulted in a total of 2,115 SNP markers for RILs-SSD and 2,354 markers for RILs-Bulk (1,272 SilicoDArTs and 1,082 SNPs) (Table 4). The total distance obtained for the map constructed in the RILs-SSD population was 1,572 cM, while for the RILs-Bulk it was 1,570 cM (Table 5).

Table 4. Filters used to obtain the set of SNPs and SilicoDArTs markers for the RILs analyses.

	RILs-SSD	RILs-Bulk
Initial number of markers	11,279	35,976
Filter 1: elimination of duplicate markers	4,914	5,482
Filter 2: elimination of monomorphic markers; missing data; heterozygous loci and distorted loci	768	28,139
Total markers after filters 1 and 2	5,597	2,354
Filter 3: Selection of markers every 50,000 bp	2,115	-
Number of markers for molecular analysis	2,115	2,354

Table 5. Distribution of markers on the chromosomes of RILs-SSD (SNPs) and RILs-Bulk (SNPs and SilicoDArTs).

	RILs-SSD			RILs-Bulk			
Chr	Total Markers	Total Distance (Mbp)	Total Distance (cM)	Total Markers SilicoDArTs	Total Markers SNPs	Total Distance (Mbp)	Total Distance (cM)
1	135	43	203	173	105	42	203
2	68	34	169	97	94	34	169
3	40	35	190	47	59	35	189
4	274	34	130	137	107	34	130
5	63	27	120	83	81	27	120
6	249	27	114	99	112	29	126
7	327	29	155	59	75	29	155
8	309	27	115	108	98	28	117
9	229	22	92	69	98	22	92
10	286	23	92	82	73	20	75
11	73	28	93	174	98	28	94
12	62	27	99	144	82	27	100
Total	2,115	356	1,572	1,272	1,082	355	1,570

Chr: chromosome; Mbp: megabase pairs; cM: centimorgan.

The linkage disequilibrium (LD) decay pattern was estimated in relation to the distance between markers for each chromosome, considering the 158 RILs from each method separately. Considering all markers, the distribution of r^2 showed a decay with the increase in physical distance between markers, presenting a high LD for both RILs-SSD and RILsBulk. On average, the LD for RILs-SSD did not decay quickly with increasing physical distance from the markers, with the average decay point being 11.58 Mb, while for RILs-Bulk RILs the decay was faster, at 0.84 Mb.

The genotyping data allowed us to estimate the genetic distance between the RILs, and according to the Smouse & Peakall coefficient, the average distance between the RILs-SSD lineages was 17,664.85, while that between the RILs-Bulk was 7,304.82, that is, the first group of RILs was almost two and a half times more variable than the second group. The expected heterozygosity (He) of these data sets also showed a higher value in the RILs-SSD group compared to the RILs-Bulk group (0.49 and 0.25, respectively).

QTL analysis

QTL analysis by multiple mapping identified six QTLs for productivity (minimum LOD of 3.0), with three QTLs identified for SSD-RILs and three QTLs for RILs-Bulk. Considering the SSD-RILs, two QTLs for GY were found for the second year of evaluation, located on chromosomes 6 and 9 (qGYLD6.1 and qGYLD9), with explained phenotypic variation values of 23.56% and 7.45%, respectively, and in the joint analysis, a QTL was identified, also located on chromosome 6 (qGYLD6.2) with an explained phenotypic variation of 9.45% (Table 6).

For SSD-RILs, considering the GY character, the QTL peak was identified by SNP_6_9538767 (in the 2018 experiment), located on chromosome 6, and is within a haplotypic block of 329 Kb, and is located between the genes LOC_Os06g16590 and LOC_Os06g16600. Considering the data from the joint analysis, the SNP_6_17931455 marker is in a 100 Kb block, and is positioned in an intergenic region. Another QTL related to productivity was identified on chromosome 9, and this QTL had as its peak of significance the marker Dart_9_22544728, which was located in the gene LOC_Os09g39240, and is in a haplotypic block with an extension of 436 Kb.

Considering the RILs-Bulk, two QTLs were located on chromosomes 6 and 9 (qGYLD6 and qGYLD9) in the 2018 experiment, with phenotypic variation of 21.65% and 3.71%, respectively. In the joint analysis, a QTL

Table 6. Quantitative Trait Loci (QTLs) for the Grain Yield trait identified in RIL populations, considering the experiment in 2018 and joint analysis (experiments in 2017 and 2018).

	QTL	Chr	Marker Interval	Distance (cM)	Block (LD)	Position (cM)	Significant Marker	LOD	% PV	ADD	SNP Alleles
2018	qGYLD6.1	6	SNP_6_7843142 to SNP_6_11219760	46 to 55	SNP_6_9563117 to SNP_6_9114303 (329 Kb)	53	SNP_6_9538767	9.38	23,56	389.51	C/T
2018	qGYLD9	9	SNP_9_15081610 to SNP_9_22925474	46 to 92	SNP_9_22489199 to SNP_9_22925474 (436 Kb)	91.7	SNP_9_22544728	3.27	7,45	205.76	A/G
Joint	qGYLD6.2	6	SNP_6_17061727 to SNP_6_22527564	68 to 90	SNP_6_17931455 to SNP_6_25390861 (100 Kb)	74	SNP_6_17931455	3.14	9,45	132.43	G/A
2018	qGYLD6	6	Dart_6_1855213 to Dart_6_2172757	5.08 to 7.43	Dart_6_18552013 to Dart_6_1901482 (46 Kb)	5.33	Dart_6_1901482	6.61	21,65	199.55	SilicoDArT
2018	qGYLD9	9	c9.loc31 to SNP_9_22925474	31.77 to 92.37	SNP_9_12730461 to SNP_9_12730961 (0 Kb)	32.77	SNP_9_12730961	3.71	11,39	144.26	C/T
Joint	qGYLD7	7	Dart_7_1138309 to Dart_7_29645766	10.74 to 155.56	Dart_7_28888504 (49 Kb)	150	Dart_7_28888504	3.23	12,90	- 70.15	SilicoDArT

Chr: chromosome; Position (cM) = QTL position on the genetic map in cM; Block (LD) = Block in linkage disequilibrium; %PV: percentage of explained phenotypical variation; ADD= Additive Effect (negative sign means allele inherited from parent 'IRAT'.

was identified on chromosome 7 (qGYLD7) with phenotypic variation of 12.90% (Table 6). With the exception of QTL qGYLD7, all other QTLs had the favorable allele coming from the parent Epagri 108. On chromosome 6, the peak of QTL significance was at Dart_6_1901782, and this marker is located in the region close to LOC_Os06g04450 within a block of 46 Kb. In the QTL identified on chromosome 7, the significant peak of significance was at Dart_7_28888504 and is located in an intergenic region. In the QTL of chromosome 9, the marker SNP_9_12730961 is positioned in the LOC_Os09g21110 gene.

Discussion

Phenotypic Analysis

The experimental coefficients of variation of this work indicate that the two years of evaluation had good to moderate experimental precision, while selective precision was moderate to high, according to Resende and Duarte (2007). The crossing between the two parents allowed good genetic complementarity, which provided the RILs with higher grain yield than their parents, resulting from transgressive segregation. The RILs-Bulk group stood out with the highest productivity averages in both years and in the joint analysis, while the checks presented lower averages than the Bulk and SSD groups. These results reflected the high and significant combinatorial capacity of the Epagri 108 and Irat 122 parents previously observed by Ramos et al. (2019).

The combining ability reflects the effect of gene additivity and is of fundamental importance in selection in autogamous plants (Bernardo et al., 2010). Kambar et al. (2011) found a similar result in rice, and Silva et al. (2013) in common bean, where the Bulk method was superior to the SSD method in obtaining inbred lines with higher average grain productivity. This indicates that the most productive plants left the greatest number of seeds and with each generation of population advancement, these plants had greater representation. This was also reflect-

ed in the smaller genetic distance between RILs-Bulk in relation to RILs-SSD obtained by SNP markers, and greater genetic variance and heritability in RILs-SSD than in SSD-Bulk, as observed by Miladinovic et al. (2011). According to Funada et al. (2013), through simulation, in the SSD method, 82% of plants in F_2 are represented in the F_6 generation. On the other hand, Kervella and Fouillou (1992), also using simulation, showed that in the Bulk method, only 67% of the plants in F_2 are represented in F_5, that is, indicating losses due to sampling of the method.

Genotypic Analysis

The molecular markers were positioned on the 12 chromosomes and ensured broad distribution throughout the genome of both RIL populations. Any factor that alters allele frequencies can interfere with the LD dynamics observed in the RIL populations of the study, such as natural or artificial selection, sampling processes, among others (Schaper et al., 2012).

In rice, linkage disequilibrium values are quite variable, for example, in inbred lines, the LD varies from 100-500 Kbp, but these values are not common and result from the specific breeding processes to which the population was subjected (Liu et al., 2013). Chen et al. (2013) observed LD in *O. sativa* Japonica subspecies from temperate regions being the one with the highest LD (> 500 Kb), followed by *O. sativa* Japonica from tropical regions (approximately 150 Kb) and *O. sativa* Indica (approximately 75 Kb).

The linkage map was constructed by 2,115 SNP markers for RILs-SSD and 2,354 markers for RILs-Bulk (SNPs and DArTs). According to Shabir et al. (2017), genetic mapping of a suitable number of markers, and a progeny with 50 to 250 individuals, provides a wide marker distribution across the rice genome, and increasing marker density greatly improves map resolution and accuracy for QTL analysis (Yonemaru et al., 2014). The map obtained in the study has 1,572 cM for RILs-SSD and 1,570 cM for RILs-Bulk, and whose distances are relatively greater in relation to the map by Ma et al. (2016) in rice, which presented 1,070 cM.

In the present work, six QTLs for productivity were identified (three for RILs-SSD and three for RILs-Bulk). Although the same QTLs were not found in the different years evaluated, the LOD values were high enough to determine that the results were quite consistent in each location, which makes the use of the QTL for assisted selection only for the environment where it was identified. According to Cobb et al. (2019), it is not possible to transfer the results of QTL analysis to other environments, years or other populations, due to the quantitative inheritance of traits such as productivity.

Silva et al. (2022), in a QTL analysis using different RILs-SSD, but derived from the same crossing Epagri 108 × Irat 122, and evaluated in the same location (Goianira, GO), but in another cropping season, identified the same SNP associated with the productivity QTL of chromosome 9 (qGYLD9; SNP_9_22544728). This result is very interesting, as two independent experiments, carried out in different years, and with different SNP data sets, identified the same SNP related to the productivity trait. This SNP is a candidate to be studied in detail for the development of a marker for assisted selection for productivity. SNP_9_22544728 (qGYLD9), which has a modifier effect, is located in an exon of the gene LOC_Os09g39240, and is translated into a protein with a function that has not yet been determined. The SNPs that identified the QTLs qGYLD6.1 and qGYLD6.2 were not located in the rice genes.

In RILs-Bulk, two QTLs were identified for the GY trait in the 2nd year. qGYLD6 was identified on chromosome 6, where the significant SNP was positioned within the LOC_Os06g04450 gene with a putative function as a transport protein of the Sec1 family, whose function has not yet been determined in plants. The qGYLD9 QTL on chromosome 9 is located in the LOC_Os09g21110 gene, responsible for the synthesis of the leucyl-tR-

NA synthetase enzyme, and although already identified in Arabidopsis, corn, sorghum and grapes, its function has been more studied in humans, and related to growth regulation cellular via the TOR complex route (Son et al., 2019). The SNP variants that identified the genes LOC_Os06g04450 (QTL qGYLD6) and LOC_Os09g21110 (QTL qGYLD9) have a modifier impact and are located in a 3' UTR region. The SNP that identified the QTL qGYLD7 is located in an intergenic region.

Two QTLs (qGYLD6.2 and qGYLD7) were detected only by joint analysis, where the productivity values are the result of the average of the genotypic values of the two years of experiment. The phenotypic contrasts between the RILs are determined by the average value of the years, finding a distinct value, resulting from the G x E interaction, which then generate a data matrix that would be equivalent to a third population. However, the joint analysis data, despite being an estimate, may reflect the identification of more stable QTL loci. According to Wang et al. (2019), joint analyzes across multiple years and locations can be used to determine QTL stability and estimate the interaction between the additive QTL and the environment.

The choice of the most efficient breeding method for managing segregating populations depends on variables such as the heritability of the trait, the financial resources available for each breeding program, the skill of breeders, the trait to be improved and the parents used (Ramalho et al., 2012). Of the limitations that a breeding program faces, the cost factor generated for experimentation and the labor required to conduct field experiments are the most impactful. Both the SSD and Bulk methods were efficient in obtaining high grain yield RILs and in the ability to identify QTLs. However, the Bulk method was more practical to handle and cheaper to obtain RILs compared to SSD.

Conclusions

1) The Bulk method obtained RILs with less genetic distance between them, but presented higher average grain yield compared to RILs derived from SSD.

2) Both RILs derived from Bulk and SSD showed sufficient phenotypic contrast to identify QTLs for the grain yield trait.

3) Considering the higher average grain yield of RILs-Bulk and the identification of RIL-SSD with higher GY, it was not possible to choose the most efficient method. However, the Bulk method is easy to manage seeds over generations, a point that may be important in order to reduce the cost of population development.

References

BARRETT, J.C.; FRY, B.; MALLER, J.; DALY, M.J. 2005. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics**, 21:263-265.

BERNARDO, R. 2010. **Breeding for quantitative traits in plants**. Stemma Press, Woodbury.

CHEN, H.; HE, H.; ZHOU, F.; YU, H.; DENG, X. W. Development of genomics-based genotyping platforms and their applications in rice breeding. **Current Opinion in Plant Biology**, 16:247-254, 2013.

COBB, J.N.; BISWAS, P.S.; PLATTEN, J.D. 2019. Back to the future: revisiting MAS as a tool for modern plant breeding. **Theoretical and Applied Genetics**, 132(3):647-667.

DESROUSSEAUX, D.; SANDRON. F.; SIBERCHICOT, A.; CIERCO-AYROLLES, C.; MANGIN, B. LDcorSV: Linkage disequilibrium corrected by the structure and the relatedness. R package version 1.3.1., 2013.

DIXIT, S.; SINGH, M.U.; ABBAI, R.; RAM, T.; SINGH, A. P.; VIRK, P. S.; KUMAR, A. 2019. Identification of genomic regions responsible for high iron and zinc content in rice. **Scientific Reports**, 9:8136.

FUNADA, M.; HELMS, T.C.; HAMMOND, J.J.; HOSSAIN, K.; DOETKOTT, C. 2013. Single-seed descent, single-pod, and Bulk sampling methods for soybean. **Euphytica**, 192:217–226.

GUO, L.; GAO, Z.; QIAN, Q. 2014. Application of resequencing to rice genomics, functional genomics and evolutionary analysis. **Rice**, 7(1):4.

HADDAD, N.I.; MUEHLBAUER, F.J. 1981. Comparison of random bulk population and single-seed-descent methods for lentil breeding. **Euphytica**, 30:643-651.

HILL, W. G.; WEIR, B. S. 1988. Variances and covariances of squared linkage disequilibria in finite populations. **Theoretical Population Biology**, 33:54–78.

KANBAR, A.; KONDO, K.; SHASHIDHAR, H. E. 2011. Comparative efficiency of pedigree, modified Bulk and single seed descent breeding methods of selection for developing high-yielding lines in rice (*Oryza sativa* L.) under aerobic condition. **Electronic Journal of Plant Breeding**, 2(2):184-193.

KERVELLA, J.; FOUILLOUX, G. 1992. A theoretical study of the Bulk breeding method I. Importance and consequences of losses due to sampling. **Euphytica**, 60:185-195.

LIU, T.; YU, T.; XING, Y. 2013. Identification and validation of a yield-enhancing QTL cluster in rice (*Oryza sativa* L.). **Euphytica**, 192(1):145-153.

Lorieux, M. 2012. MapDisto: fast and efficient computation of genetic linkage maps. **Molecular Breeding**, 30:1231–1235.

MA, X.; FU, Y.; ZHAO, X.; JIANG, L.; ZHU, Z.; GU, P.; XU, W.; SU, Z.; SUN, C.; TAN, L. 2016. Genomic structure analysis of a set of *Oryza nivara* introgression lines and identification of yield associated QTLs using whole genome resequencing. **Scientific Reports**, 6:27425.

McCOUCH, S. R.; CHEN, X.; PANAUD, O.; TEMNYKH, S.; XU, Y.; CHO, Y. G.; HUANG, N.; ISHII, T.; BLAIR, M. 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. **Plant Molecular Biology**, 35(1-2), 89-99.

MENDES, M.P.; RAMALHO, M.A.P.; ABREU, A.F.B. 2011. Strategies in identifying individuals in a segregant population of common bean and implications of genotype x environment interaction in the success of selection. **Genetics and Molecular Research**, 11: 872-880.

MILADINOVIĆ, J.; BURTON, J.W.; BALEŠEVIĆ, S.; MILADINOVIĆ, D.; DJORDJEVIĆ, V.; DJUKIC, V. 2011. Soybean breeding: comparison of the efficiency of different selection methods. **Turkish Journal of Agriculture and Forestry**, 35:469-480.

PEAKALL, R.; SMOUSE, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. **Bioinformatics**, 28:2537-2539.

QI, P.; LIN, Y.-S.; SONG, X.-J.; SHEN, J.-B.; HUANG, W.; SHAN, J.-X.; ZHU, M.-Z.; JIANG, L.; GAO, J.-P.; LIN, H.-X. 2012. The novel quantitative trait locus GL3.1 controls rice grain size and yield by regulating Cyclin-T1;3. **Cell Research**, 22 (12):1666–1680.

R CORE TEAM. 2022. **R: A language and environment for statistical computing**. Available from URL. http://www.r-project.org/.

RAMALHO, M.A.P.; ABREU, A.F.B.; SANTOS, J.B.; NUNES, J.A.R. 2012. **Aplicações da genética quantitativa no melhoramento de plantas autógamas**. UFLA, Lavras. 522p.

RAMOS, M. R. F.; MENDONÇA, J. A.; VIANELLO, R. P.; MORAIS JÚNIOR, O. P.; COLOMBARI, J. M.; BORBA, T. C. O.; CASTRO, A. P.; BRONDANI, C. Heterosis And Combining Ability For Grain Yield And Earliness In Accessions Of a Rice Core Collection. **Functional Plant Breeding Journal**, 1(1):1-13.

RESENDE, M. D. V.; DUARTE, J. B. 2007. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, 37(3):182-194.

SCHABIR, G.; ASLAM, K.; KHAN, A.; SHAHID, M.; MANZOOR, H. et al. 2017. Rice molecular markers and genetic mapping: current status and prospects. **Journal of Integrative Agriculture**, 16(9):1879-1891.

SCHAPER, E.; ERIKSSON, A.; RAFAJLOVIC, M.; SAGITOV, S.; MEHLIG, B. 2012. Linkage disequilibrium under recurrent bottlenecks. **Genetics**, 190(1):217-229.

SILVA, A.C.; MELO, P.G.S.; MELO, L.C.; BASSINELO, P.Z.; PEREIRA, H.S. 2013. Eficiência de métodos de melhoramento para teor de fibra e produtividade de grãos em progênies de feijoeiro comum. **Bragantia**, 72(4):326-331.

SILVA, D.A.R.; MENDONÇA, J.A.; CORDEIRO, A.C.C.; MAGALHÃES JÚNIOR, A.M. de; VIANELLO, R.P.; BRONDANI, C. 2022. Identification of stable quantitative trait loci for grain yield in rice. **Pesquisa Agropecuária Brasileira**, 57:e02812.

SMOUSE, P.E.; PEAKALL, R. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. **Heredity**, 82:561-573, 1999.

SON, K.; YOU, J.; YOON, M. et al. Nontranslational function of leucyl-tRNA synthetase regulates myogenic differentiation and skeletal muscle regeneration. 2019. **Journal of Clinical Investigation**, 129(5):2088-2093.

WANG, L.; YANG, X.; CUI, S.; MU, G.; SUN, X.; LIU, L.; LI, Z. 2019. QTL mapping and QTL × environment interaction analysis of multi-seed pod in cultivated peanut (*Arachis hypogaea* L.). **The Crop Journal**, 7:249–260.

XING, Y.; ZHANG, Q. 2010. Genetic and Molecular Bases of Rice Yield. **Annual Review of Plant Biology**, 61:421–442.

YANG, J.; BENYAMIN, B.; McEVOY, B. P.; GORDON, S.; HENDERS, A.K.; NYHOLT, D.R.; MADDEN, P.A.; HEATH, A.C.; MARTIN, N.G.; MONTGOMERY, G.W.; GODDARD, M.E.; VISSCHER, P.M. 2010. Common SNPs explain a large proportion of the heritability for human height. **Nature Genetics**, 42:565–569.

YANG, J.; LEE, S. H.; GODDARD, M. E.; VISSCHER, P. M. 2011. GCTA: A tool for genome-wide complex trait analysis. **The American Journal of Human Genetics**, 88:76–82.

YONEMARU, J.; MIZOBUCHI, R.; KATO, H.; YAMAMOTO, T.; YAMAMOTO, E.; MATSUBARA, K.; HIRABAYASHI, H.; TAKEUCHI, Y.; TSUNEMATSU, H.; ISHII, T.; OHTA, H.; MAEDA, H.; EBANA, K.; YANA, M. 2014. Genomic regions involved in yield potential detected by genome-wide association analysis in Japanese high-yielding rice cultivars. **BMC Genomics**, 15:346.