Original Research Article

Genome wide association studies to dissect genetic factors conferring sheath blight resistance in Rice (*Oryza sativa* L.,)

ABSTRACT

Rice Sheath blight (ShB) is one of the most serious fungal diseases caused by *Rhizoctonia solani*. Breeding for sheath blight resistance has been ineffective exercise so far, mainly because of lack of good number of reliable sources of resistance in rice germplasm. In this context our studies indicated that the lines Tetep, Jasmine 85 and MTU 9992 confer resistant to moderately resistant reaction against the pathogen. The current investigation was carried out to dissect the genetic factors governing resistance to sheath blight through genome wide association study (GWAS) from the mapping populations developed by design where in, each of the resistant parents were crossed to three to four highly susceptible parents to generate eleven populations (Jasmine 85XTN1, Jasmine 85XSwarnaSub1, Jasmine 85XII32B, Jasmine 85XIR54, TetepXTN1, TetepXSwarnaSub1, TetepXII32B, TetepXIR54, MTU 9992XTN1, MTU 9992XII32B and MTU 9992XIRBB4). A total of 1545 Recombinant inbred lines (RILs) derived from eleven crosses were used for the study. During rainy 2020 the F7 RILs were screened for their reaction to Sheath blight in two hot spot locations. The genotyping was done with Illumina platform having 6564 SNP markers. Genome wide association study was done with two models Generalized Linear Model (GLM) and Mixed Linear Model (MLM). Results clearly indicate the superiority of MLM over GLM in correcting the population structure. With MLM model, in Jasmine 85 half-sib populations with 565 RILs analyzed, five QTLs (Quantitative Trait Loci) were detected on Chr1, Chr3, Chr9, Chr10 and Chr11 with $-\log_{10}$ (P-Value) more than 3. In TETEP half-sib populations with 714 RILs examined, seven QTLs were observed on Chr1, Chr2, Chr5, Chr6, Chr7, Chr8, and Chr11 with $-\log_{10}$ (P-Value) more than 4. Whereas in MTU 9992 half-sib populations with 266 RILs studied, three novel QTLs were identified on Chr2, Chr6 and Chr11 with $-\log_{10}$ (P-Value) more than 3. Some of these QTLs were reported by researches earlier. In the current research, some novel QTLs were detected in Jasmine 85 (Chr10) and Tetep (Chr2, Chr5 and

UNDER PEER REVIEW

Chr6) apart from three new QTLs discovered in MTU 9992. The results facilitated to have better understanding of the genetic basis for sheath blight resistance in rice. Pyramiding all the QTL identified so far into a susceptible varieties is complicated affair as resistance is governed by not only several large effect QTLs but also medium to small effect QTLs as well, hence genomic selection approach could be rewarding for breeding for sheath blight resistance.

Keywords: *Rice, sheath blight, SNPs, GWAS, LD Mapping, Association Mapping*

1. INTRODUCTION

Rice (*Oryza sativa* L.) feeds more than half of the world's population and genetic improvement of this food crop can serve as a major component of sustainable food production. Rice sheath blight (ShB) is one of the most devastating fungal diseases of rice, causing significant yield losses in many rice-growing regions of the world. This disease has become popular recently because of intensification of rice-cropping systems with the development of new short stature, high tillering, high yielding cultivars, high plant densities, and an increase in nitrogen fertilization, these morphological and microenvironment situations are very much congenial for the growth and multiplication of the sheath blight fungus, in India it's prevalence is mainly confined to coastal places of India where farmers grow very high yielding varieties and hot humid climate adds to that. These factors promote disease spread by providing a favorable microclimate for the disease agent due to a dense leaf canopy with an increased leaf-to-leaf and leaf-to-sheath contact (Banniza *et al*., 2007).

The necrotrophic Sheath Blight pathogen possess a broad range of hosts, there are few germplasm lines in Rice which are known to show resistant reaction against this pathogen, most of the breeders are focused on harnessing these resistant sources to breed cultivars which are resistant to tolerant for this disease. Because of lack of authentic and reliable sources of resistance, breeding for sheath blight has been challenging in Rice (Jia *et al*. 2009; Zuo *et al*. 2010; Srinivasachary, Willocquet and Savary 2011). There have been many studies which reported on the existence of sources with diverse levels of

UNDER PEER REVIEW

resistance in Xiangzaoxian 19 (Che *et al*. 2003), WSS2 (Sato et al. 2004), Teqing (Li *et al.* 1995; Pinson *et al*. 2005), Pecos (Sharma *et al*. 2009), Tetep (Sha and Zhu, 1989; Channamallikarjuna *et al*. 2010), Jasmine 85 (Pan *et al*. 1999; Zou *et al*. 2000; Liu *et al*. 2009), Minghui63 (Han *et al.* 2002), and wild rices O. rufipogon, O. nivara etc. (Ram *et al.* 2008; Eizenga *et al.* 2013).

Upon intensive study it's believed to be controlled by many genomic regions dispersed across the genome (Sha and Zhu 1989; Li *et al*. 1995; Pinson et al. 2005; Zuo *et al*. 2013). It is widely believed that quantitative nature of resistance could be advantageous for evolving varieties with durable/horizontal resistance (Young, 1996; Poland *et al*. 2009).

As of now, around 50 ShB resistance quantitative trait loci (ShBR QTLs) have been mapped to all the 12 rice chromosomes (Jia *et al*. 2009; Zuo *et al*. 2010; Xu *et al*. 2011 and Wang *et al.* 2012) because of advancement in genotyping technology and availability of genotypic information at cheaper price. The current research was undertaken to understand genetic basis and identify novel genomic regions governing sheath blight resistance in Rice. To unravel the new QTLs conferring resistance to sheath blight GWAS or Association mapping (AM) or Linkage Disequilibrium (LD) mapping was conducted using RILs developed from three sources of resistance. GWAS is a great tool for identification of the genomic regions controlling phenotype of interest, it exploits historical recombination events to trace and map trait variations. A major hurdle in AM is controlling false positives and false negatives that can arise from population structure and family relatedness. False positives and negatives can often be controlled by incorporating covariates for structure and kinship in mixed linear models (MLM).

2. MATERIAL AND METHODS

2.1 Parent material and phenotyping of F_7 RILs for ShB

A total of 250 germplasm lines were screened for identification of lines which were resistant and susceptible for Sheath blight. Half sib crosses were created by crossing each resistant lines with three to four agronomically superior susceptible lines to develop RIL populations involving Jasmine 85, Tetep & MTU 9992 as resistant parents and TN1, Swarna Sub1, II32B, IR54 & IRBB4 as susceptible parents. The RILs were generated by following single seed descent method (SSD) at Rapid Generation Advancement/ Speed breeding facility of Pioneer Hi-Bred Pvt. Ltd. Research Centre at Tunkikalsa village, Medak district,

UNDER PEER REVIEW

Telangana. The eleven crosses used for the study were grouped into three half-sib hubs with RILs ranged from 50 to 241 in each family (Table 2). Half-sib hub of Jasmine 85 had 565 RILs (Jasmine 85XTN1, Jasmine 85XSwarnaSub1, Jasmine 85XII32B and Jasmine 85XIR54), half-sib hub of Tetep possessed 714 RILs (TetepXTN1, TetepXSwarnaSub1, TetepXII32B and TetepXIR54) and half-sib hub of MTU9992 had 266 RILs (MTU 9992XTN1, MTU 9992XII32B and MTU 9992XIRBB4). The total of 1545 RILs derived from these eleven crosses were phenotyped for sheath bight reaction in two hot spot locations (Seethanagaram and Draksharam) of East Godavari District of Andhra Pradesh state, India(Latitude $16^0$08' N and Longitude $81^0$08' E, Latitude 17°10'N and Longitude 81°41' E).

The experiment consisted of F_7 progenies along with parental lines were planted in randomized complete design with two replications. Row length of 1.2 meter with row-to-row distance 15 cm and plant to plant distance 10 cm was considered to ensure dense population which is congenial for the development of disease. TN1 was used as susceptible check and was sown after every two rows as well as all along the border to increase the disease pressure so as to serve as spreader rows. In the present study, the virulent local East Godavari isolate of rice sheath blight pathogen was utilized for disease screening. Before the inoculation, the fungus was cultivated in potato dextrose agar medium at optimal temperature for 3–4 days, followed by transferring of disc of medium with mycelia for multiplication. To ensure stringent screening for better disease development, artificial inoculation was done by spraying the mycelia uniformly at the base of plant at maximum tillering stage. The data was recorded at peak milking stage to dough stage by visualizing the relative lesion length to height (%) using 1-9 scale based on development of lesion from the lower to upper part of plant on a scale from 1 (Resistant) to 9 (Susceptible) thereby getting total of six phenotypic classes, where score 1: no infection, score 2: 1-20%, score 3: 21-30%, score 5: 31-45%, score 7: 46-65%, score 9: 66-100%.

2.2 SNP genotyping

All the RILs used for the study were genotyped using Infinium marker platform which is a fixed plex comprising of 6564 markers, the genotyping was done at marker technology lab of Pioneer Hi-Bred International Limited at Johnston, Iowa State, United States of America.

UNDER PEER REVIEW

2.3 Description of association mapping (AM) models

The statistical analysis was done with "TASSEL" application. TASSEL also known as **Trait Analysis by aSSociation, Evolution and Linkage** is a powerful statistical software to conduct association mapping such as General Linear Model (GLM) and Mixed Linear Model (MLM).

In the current study, two models were used for statistical analysis, (i) general linear model (GLM) with PCoA (principle coordinate analysis) (Price *et al.,* 2006), (ii) mixed linear model (MLM) with PCoA + K (Kinship matrix for family relatedness estimates) (Yu *et al*., 2006).

The simplest model is to directly detect the association between a phenotype (y) and markers (Si) one at a time, where i=1 to m, and m is number of markers. In GLM, in order to reduce spurious associations (false positives) while performing association mapping (AM) consideration of population structure (Q) as a cofactor helps in accounting residuals (e) partially and also adjusts some effect that does not belong to the testing markers. The mixed linear model (MLM) applies the same principle by adding individuals' genetic effects as random cofactor effects with variance structure defined by the kinship (K) among individuals. In both Q and Q+K models, Q and K stay the same. Because of inclusion of additional cofactor family relatedness in the model in case of MLM, both false positives and false negatives are taken care.

GLM model equation: $y = S_i + Q/PCA/PCoA + e$

MLM model equation: $y = S_{i+} Q/PCA/PcoA + K + e$

The analysis was done with both GLM and MLM for all three half-sib hubs separately involving Jasmine 85, Tetep and MTU 9992 to systematically trace the genomic regions governing the phenotype under study.

3. RESULTS AND DISCUSSION

The frequency distribution of 1545 F_7 progenies evaluated showed continuous variation across all half-sib population studied (Figure 1, 2 and 3). The genotypic analysis results were compelling because of usage of large number of markers and excellent distribution of the markers throughout the genome (Table 1), polymorphic markers between parents across population ranged from 1407 to 2849, MTU 9992/TN1 and MTU 9992/IRBB4 possessed lowest and highest number of informative markers (Table 2).

UNDER PEER REVIEW

Association mapping relies mainly on the LD between marker and QTL, the main reason for false positives in AM is Linkage disequilibrium, LD can be observed because of population structure, selection, random drift, familial relatedness. Hence it is important to separate the LD of the marker with QTL from LD due to other reasons. By inclusions of population structure and familial relatedness cofactors in the model, spurious associations were taken care. The power signal detection is determined by several factors including the heritability of trait, population structure, extent of LD in populations, size of the population, pollination mechanism of crop species (Yu *et al*., 2006).

The results of Principal co-ordinate analysis/multidimensional scaling method clearly indicated that there was enough diversity among the populations present in each half-sib hubs (Figure 4, 5 and 6). With MLM model in Jasmine 85 half-sib populations, five QTLs (Quantitative Trait Loci) were found on Chr1, Chr3, Chr9, Chr10 and Chr11 with $-\log_{10}$ (*P*-Value) more than 3, the results were similar with signals detected on Chr1, Chr3, Chr9, Chr10 and Chr11 with GLM as well (Figure 7 and 8), the signals detected were near the proximity where some of the QTLs were mapped already, Q*Rh1* (Chr1), q*SB-3 (*Chr3), q*ShB9-2 (*Chr9) and q*SB-11 (*Chr11).

In Tetep half-sib populations, seven QTLs were observed on Chr1, Chr2, Chr5, Chr6, Chr7, Chr8, and Chr11 with $-\log_{10}$ (*P*-Value) more than 4 with MLM model, whereas GLM exhibited signals on all chromosomes with $-\log_{10}$ (*P*-Value) more than 4 (Figure 9 and 10), the signals identified were near the region where QTLs were mapped by earlier researchers, q*SBR1-1* (Chr1), q*SBR7-1* (Chr7), q*SBR8-1* (Chr8) and q*SBR1-1* (Chr11). However in MTU 9992 populations, three novel QTLs were discovered on Chr2, Chr6 and Chr11 with $-\log_{10}$ (*P*-Value) more than 3 with MLM model, the results were similar with signals detected on Chr2, Chr6 and Chr11 in case of GLM (Figure 11 and 12). There have been many studies which reported sheath blight QTLs on multiple chromosomes in Jasmine 85 (Pan *et al.* 1999; Zou *et al*. 2000; Liu *et al*. 2009) and Tetep (Sha and Zhu, 1989; Channamallikarjuna *et al.* 2010).

In the current investigation, the MLM model performed better in removing spurious associations and detecting signals distributed throughout the genome with much more precision by inclusion of familial relatedness cofactor in the model. some novel QTL were detected in Jasmine 85 (Chr10), Tetep (Chr2, Chr5 and Chr6) and MTU 9992 (Chr2, Chr6 and Chr11), these have to be fine mapped and validated for

UNDER PEER REVIEW

their efficacy to use further in breeding for sheath blight resistance. However, looking into strength of signals and marker effects generated after statistical analysis clearly indicated that several loci with medium to small effects scattered across the genome did contribute to sheath blight resistance in each resistant parent which hinted that the resistance to sheath blight is governed by many genes with additive effect, this was reported by earlier researchers (Sha and Zhu 1989; Li *et al*. 1995; Pinson *et al*. 2005; Zuo *et al.* 2013).

4. CONCLUSION

The results of the current investigation facilitated to discover new regions controlling resistance and helped to have better understanding of the genetic basis for sheath blight resistance in rice. Pyramiding all the QTL identified so far into a susceptible varieties is challenging task as resistance is governed by not only several large effect QTLs but also medium to small effect QTLs as well. The inheritance of disease resistance is complex, hence genomic selection approach could be rewarding for breeding for sheath blight resistance as genomic selection considers marker effects of all loci dispersed across the genome to provide genomic estimated breeding values which can be used for selection or rejection of breeding lines with resistance to sheath blight.

ACKNOWLEDGEMENTS

I would like to thank Corteva Agriscience (Pioneer Hi-Bred Private Limited, Tunkikalsa Village, Medak, Telangana State, India) for providing all the facilities to carry out my research work. I greatly acknowledge my advisory committee members for their suggestions, support and guidance.

UNDER PEER REVIEW

Figure 1: Frequency distribution of ShB phenotypic scores for half-sib families of Jasmine 85

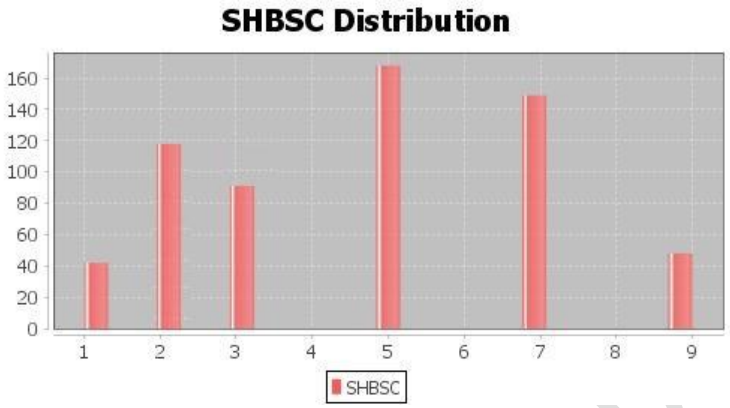

Figure 2: Frequency distribution of ShB phenotypic scores for half-sib families of Tetep

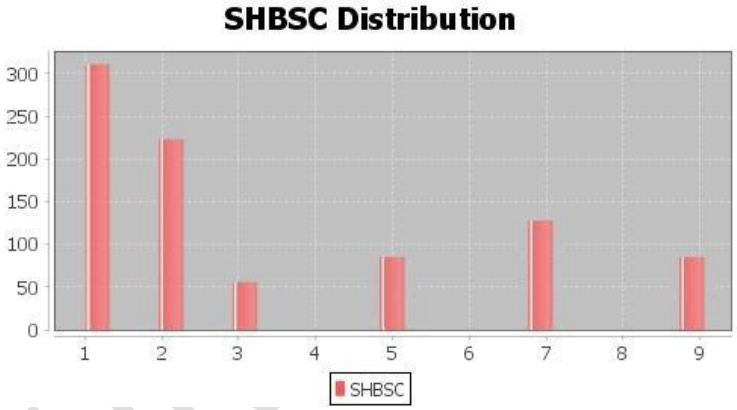

Figure 3: Frequency distribution of ShB phenotypic scores for half-sib families of MTU 9992.

UNDER PEER REVIEW

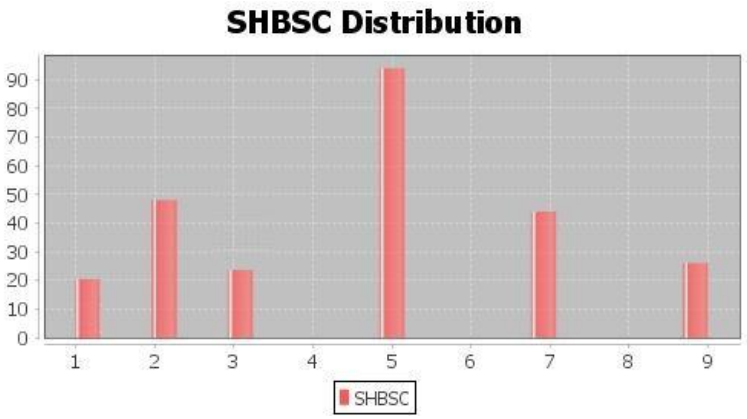

Figure 4: Depiction of analysis results of principal components (PCoA) in Jasmine 85 half-sib families

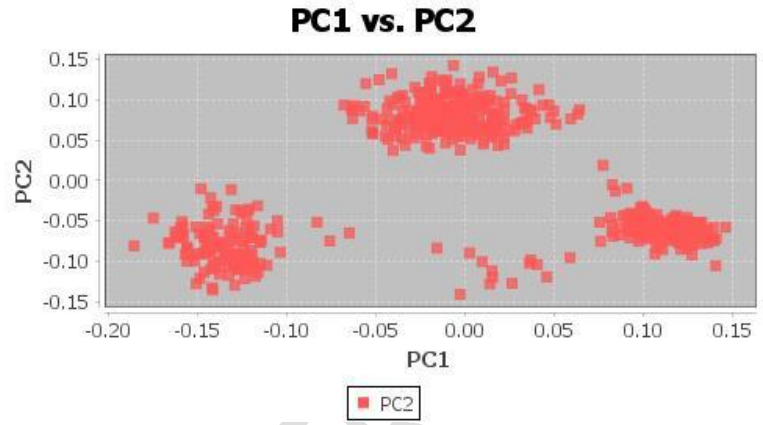

Figure 5: Depiction of analysis results of principal components (PCoA) in Tetep half-sib families

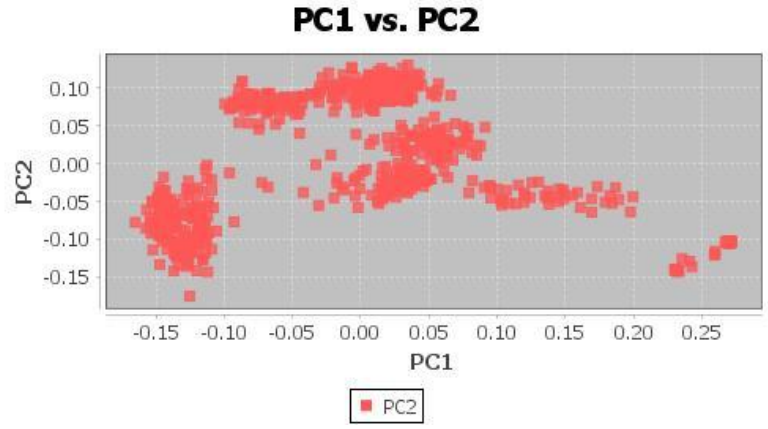

Figure 6: Depiction of analysis results of principal components (PCoA) in MTU 9992 half-sib families

UNDER PEER REVIEW

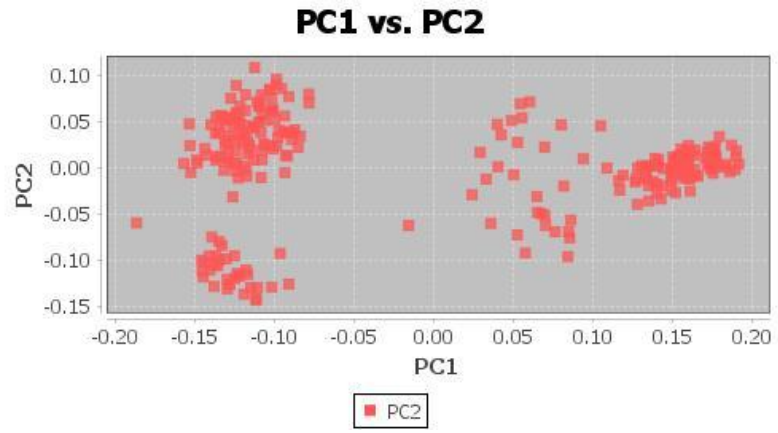

Figure 7: Manhattan plot depicting genome wide association results for sheath blight in Jasmine 85 half-sib populations using generalized linear model (**GLM**) for analysis.

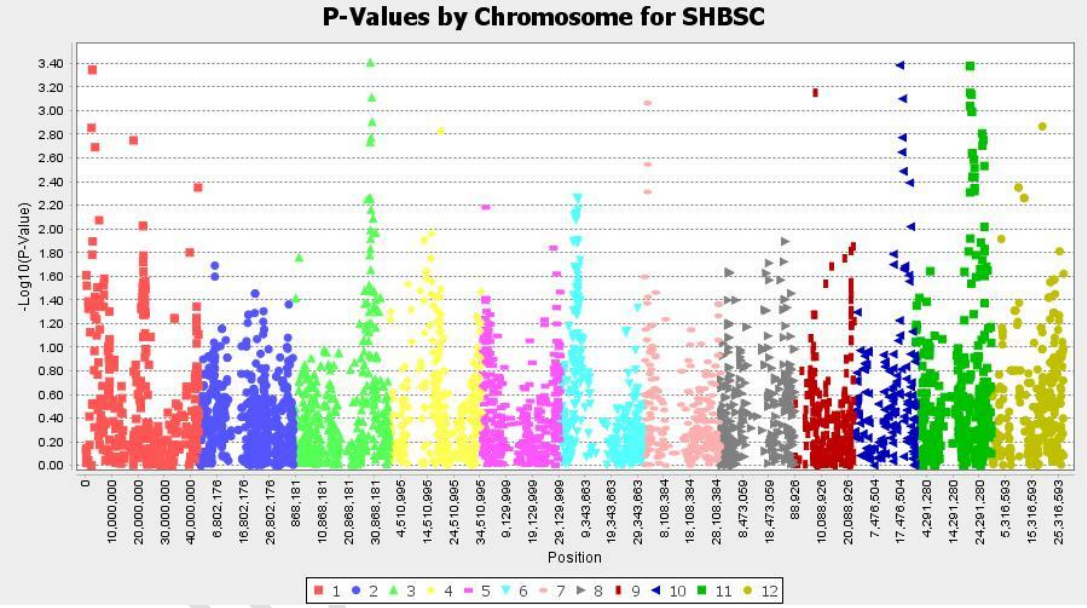

Figure 8: Manhattan plot depicting genome wide association results for sheath blight in Jasmine 85 half-sib populations using mixed linear model (**MLM**) for analysis

UNDER REVIEW

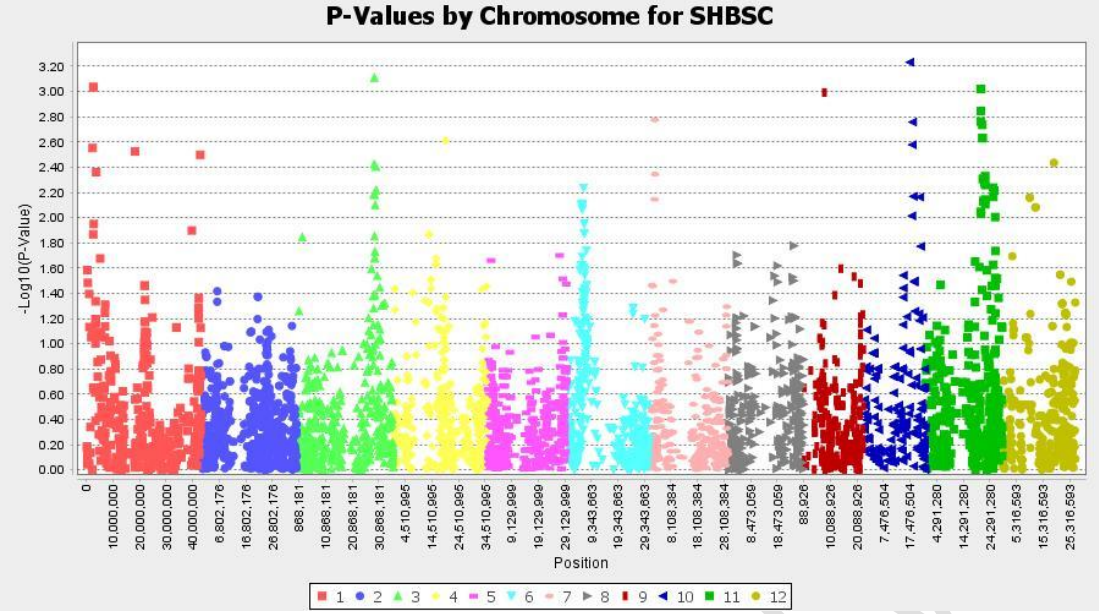

Figure 9: Manhattan plot depicting genome wide association results for sheath blight trait of Tetep half-sib populations using generalized linear model (**GLM**) for analysis.

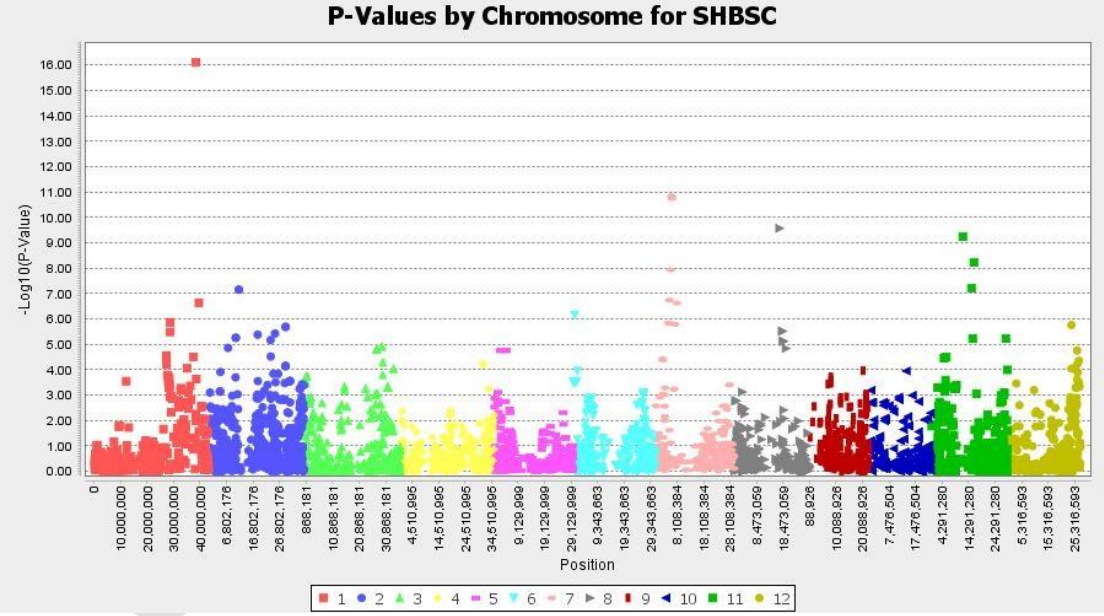

Figure 10: Manhattan plot depicting genome wide association results for sheath blight in Tetep half-sib populations using mixed linear model (**MLM**) for analysis.

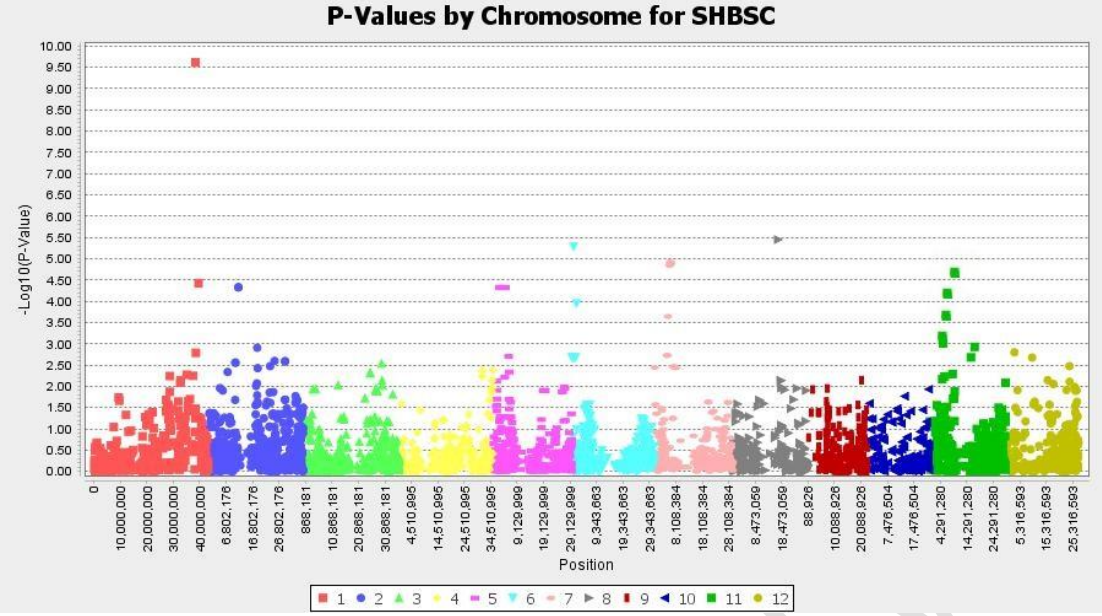

Figure 11: Manhattan plot depicting genome wide association results for sheath blight in MTU 9992 half-sib populations using generalized linear model (**GLM**) for analysis.

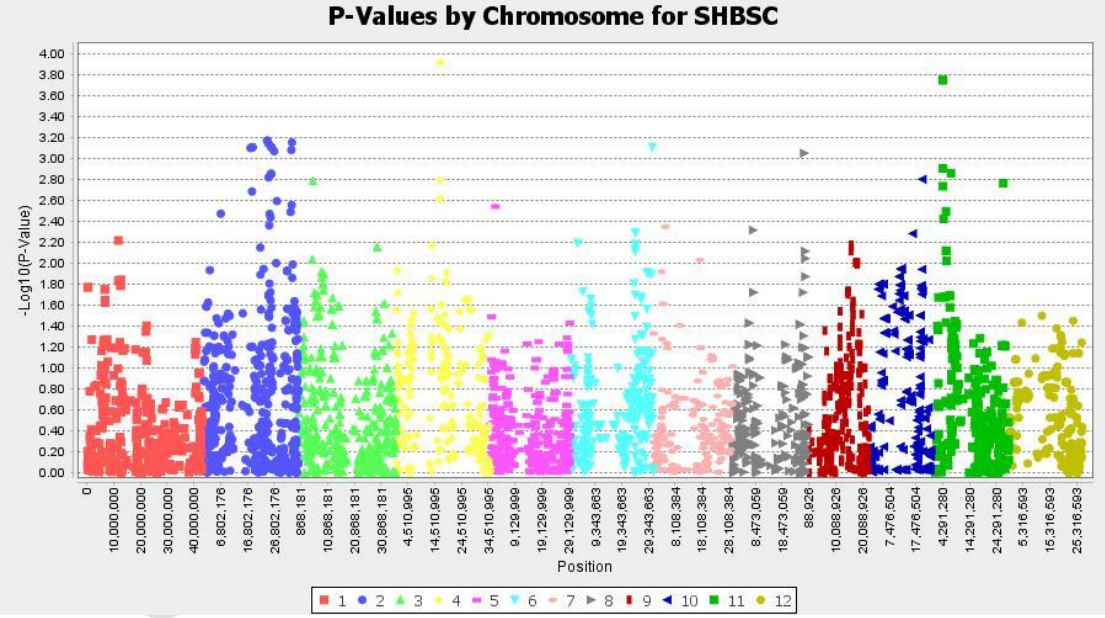

Figure 12: Manhattan plot depicting genome wide association results for sheath blight trait of MTU 9992 half-sib populations using mixed linear model (**MLM**) for analysis.

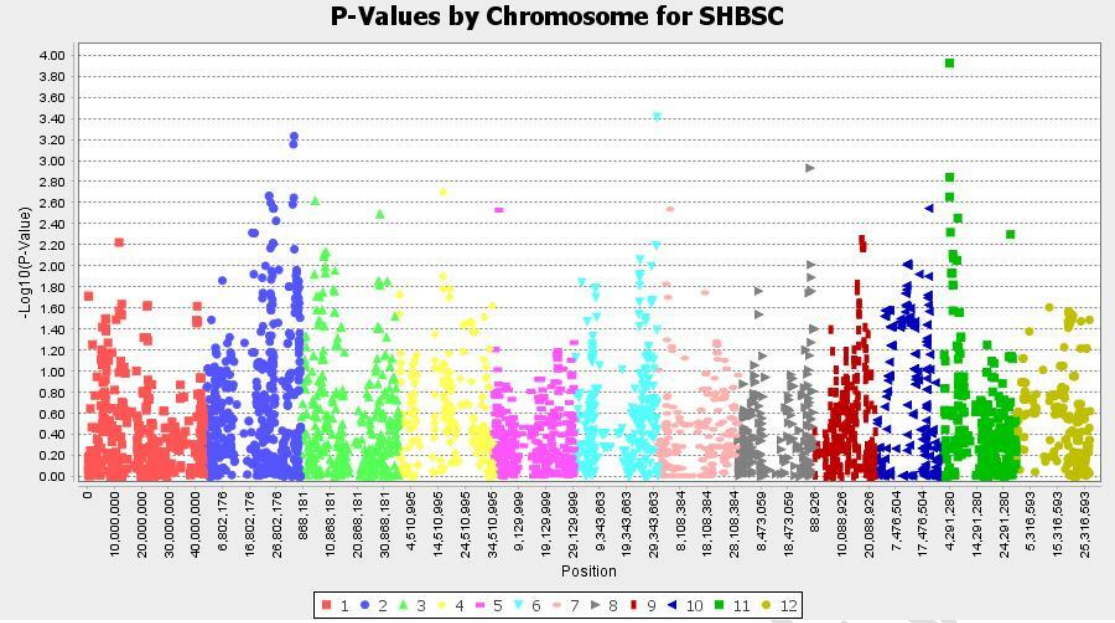

Table 1: Summary of marker data used for analysis and SNPs distribution on each chromosome

Chromosome	SNPs	Length (cM)
Ch1	639	181.8
Ch2	846	162.84
Ch3	598	164.04
Ch4	594	129.6
Ch5	583	128.58
Ch6	577	124.4
Ch7	457	118.6
Ch8	495	121.2
Ch9	427	93
Ch10	324	84.01
Ch11	541	117.9
Ch12	483	109.5
Total	**6564**	**1535.47**

Table 2: The informative markers available across the genome for each population used for analysis

Populations	No of RILs	Total Markers	Polymorphic Markers

UNDER PEER REVIEW

Jasmine 85/TN1	121	6564	2522
Jasmine 85/Swarna Sub1	139	6564	2627
Jasmine 85/II32B	144	6564	2586
Jasmine 85/IR54	161	6564	2663
Tetep/TN1	221	6564	2806
Tetep/Swarna Sub1	158	6564	2278
Tetep/II32B	241	6564	2702
Tetep/IR54	94	6564	2796
MTU 9992/TN1	50	6564	1407
MTU 9992/II32B	122	6564	2314
MTU 9992/IRBB4	94	6564	2849

UNDER PEER REVIEW

COMPETING INTERESTS DISCLAIMER:

Authors have declared that no competing interests exist. The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors.

REFERENCES

Banniza S, A. A. Sy, P. D. Bridge, S. A. Simons, and M. Holderness (2007) Characterization of Populations of Rhizoctonia solani in Paddy Rice Fields in Côte d'Ivoire. Published Online: 22 Feb 2007https://doi.org/10.1094/PHYTO.1999.89.5.414

Channamallikarjuna V, Sonah H, Prasad M, Rao GN, Singh NK, Sharma TR (2010) Identification of major quantitative trait loci qSBR11–1 for sheath blight resistance in rice. Mol Breed 25:155–166

Che KP, Zhan QC, Xing QH, He DJ, Wang B (2003) Tagging and mapping of rice sheath blight resistant gene. Theor Appl Genet 106:293–297

Eizenga GC, Prasad B, Jackson AK, Jia MH (2013) Identification of rice sheath blight and blast quantitative trait loci in two different O. sativa/O. nivara advanced backcross populations. Mol Breed 31(4):889–907

Han PY, Xing ZY, Chen XZ, Gu LS, Pan BX, Chen LX, Zhang FQ (2002) Mapping QTLs for horizontal resistance to sheath blight in an elite rice restorer line, Minghui 63. Acta Gen Sin 29(4):622–626

Jia Y, Liu GJ, Costanzo S, Lee SH, Dai YT (2009) Current progress on genetic interactions of rice with rice blast and sheath blight fungi. Front Agri in China 3:231–239

Li ZK, Pinson SRM, Marshetti MA, Stansel JW, Park WD (1995) Characterization of quantitative trait loci in cultivated rice contributing to field resistance to sheath blight(Rhizoctonia solani). Theor Appl Genet 91:374–381

Liu G, Jia Y, McClung KM, Datta A, Correll JC (2009) Mapping quantitative trait loci responsible for resistance to sheath blight in rice. Phytopath 99:1078–1084

Pan XB, Rush MC, Sha Y, Xie QJ, Linscombe SD, Stetina SR, Oard JH (1999) Major gene, nonallelic sheath blight resistance from the rice cultivars Jasmine85 and Teqing. Crop Sci 39:338–346

Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ (2009) Shades of gray: the world of quantitative disease resistance. Trends in Pl Sci 14:21–29

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., et al. (2006). Principal components analysis corrects for stratification in genomewide association studies. Nat. Genet. 38 (8), 904–909. doi: 10.1038/ng1847

Ram T, Majumdar ND, Laha GS, Ansari MM, Kar CS, Mishra B (2008) Identification of donors for sheath blight resistance in wild rice. Indian J of Gen 68:317–319

Sato HI, Deta O, Audo J, Kunihiro Y, Hirabayashi H, Iwano M, Imbe T (2004) Mapping QTLs for sheath blight resistance in the rice line WSS2. Breed Sci 54:265–271

Sha XY, Zhu LH (1989) Resistance of some rice varieties to sheath blight (ShB). Int Rice Res Newsl 15:7–8

UNDER PEER REVIEW

Sharma A, McClung AM, Pinson SRM, Kepiro JL, Shank AR, Tabien RE, Fjellstrom R (2009) Genetic mapping of sheath blight resistance QTL within tropical japonica rice cultivars. Crop Sci 49:256–264

Srinivasachary L, Willocquet L, Savary S (2011) Resistance to rice sheath blight (Rhizoctonia solani Kuhn) [teleomorph: Thanatephoruscucumeris (A.B. Frank) Donk.] disease: Current status and perspectives. Euphytica 178:1-22

Wang Y, Pinson SRM, Fjellstrom RG, Tabien RE (2012) Phenotypic gain from introgression of two QTL, qSB9-2 and qSB12-1 for rice sheath blight resistance. Mol Breeding 30:293–303

Xu Q, Yuan XP, Yu HY, Wang Y, Tang SX, Wei XH (2011) Mapping quantitative trait loci for sheath blight resistance in rice using double haploid population. Plant Breed 130:404–406

Young ND (1996) QTL mapping and quantitative disease resistance in plants. Ann Rev Phytopath 34:479–501

Yu, J., Pressoir, G., Briggs, W. H., Vroh, B. I., Yamasaki, M., Doebley, J. F., et al. (2006). A unifed mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38, 203–208. doi: 10.1038/ng1702

Zou JH, Pan XB, Hen JY, Xu JF, Lu WX, Zhu LH (2000) Mapping quantitative trait loci controlling sheath blight resistance in two rice cultivars (Oryza sativa L.). Theor Appl Genet 101:569–573

Zuo SM, Yin YJ, Zhang L, Zhang YF, Chen ZX, Pan XB (2013) Fine mapping of qSB-11LE, the QTL that confers partial resistance on rice sheath blight. Theor Appl Genet 126:1257–1272

Zuo SM, Zhang YF, Chen ZX, Chen XJ, Pan XB (2010) Current progress on genetics and breeding in resistance to rice sheath blight. Scientia Sin Vitae 40:1014–1023

UNDER PEER REVIEW

UNDER PEER REVIEW