

1 Loss of premature stop codon in the *Wall-Associated Kinase 91 (OsWAK91)* gene confers sheath
2 blight disease resistance in rice

3

4 Authors

5 ^{1,2}Noor Al-Bader albader@oregonstate.edu

6 ¹Austin Meier austin.e.meier@gmail.com

7 ^{1,2}Matthew Geniza matthew.geniza@gmail.com

8 ³Yamid Sanabria Gongora yamid.sanabria@gmail.com

9 ³James Oard joard@agcenter.lsu.edu

10 ¹Pankaj Jaiswal jaiswalp@oregonstate.edu

11

12 ¹Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, USA

13 ²Molecular and Cellular Biology Graduate Program, Oregon State University, Corvallis, OR,

14 USA

15 ³Louisiana State University Agricultural Center, Baton Rouge, LA, USA

16

17 Correspondence:

18 Pankaj Jaiswal, Ph.D.

19 jaiswalp@science.oregonstate.edu

20

21

22

23 **Abstract**

24 The genetic arms race between pathogen and host plant is a tug of war that has been ongoing for
25 millennia. The “battles” are those of disruption, restoration of signaling and information
26 transmission on a subcellular level. One such battle occurs between rice an important crop that
27 feeds 50% of the world population and the sheath blight disease (SB) caused by the fungus
28 *Rhizoctonia solani*. It results in 10–30% global yield loss annually and can reach 50% under
29 severe outbreak. Many Receptor-like kinases (RLKs) are recruited as soldiers in these battles.
30 Wall Associated Receptor Kinases (WAKs) a subfamily of receptor-like kinases have been
31 shown to play a role in fungal defense. Here we show that rice gene *OsWAK91*, present in the
32 major SB resistance QTL region on Chromosome 9 is a key component in defense against rice
33 sheath blight. An SNP mutation C/T separates susceptible variety, Cocodrie (CCDR) from the
34 resistant line MCR010277 (MCR). The resistant allele C results in the stop codon loss that
35 results in 68 amino acids longer C-terminus carrying longer protein kinase domain and
36 phosphorylation sites. Our genotype and phenotype analysis of the top 20 individuals of the
37 double haploid SB population shows a strong correlation with the SNP. The susceptible allele
38 appears as a recent introduction found in the japonica subspecies reference genome and a
39 majority of the tropical and temperate japonica lines sequenced by the 3000 rice genome project.
40 Multiple US commercial varieties with japonica background carry the susceptible allele and are
41 known for SB susceptibility. This discovery opens the possibility of introducing resistance
42 alleles into high yielding commercial varieties to reduce yield losses incurred by the disease.
43

44 **Introduction**

45 The genetic arms race between pathogen and host plant is a tug-of-war that has been
46 going on for millennia. The “battles” are those of disruption, restoration of signaling and
47 information transmission on a subcellular level. One such battle occurs between rice, an
48 important crop that feeds 50% of the global population¹ and the soil-borne pathogen *Rhizoctonia*
49 *solani* that causes the leaf sheath blight disease (SB). Leaf sheath blight is a major disease in rice
50 that negatively affects crop yield and quality². Identified by the lesions on the leaf sheath, SB
51 causes leaves and tillers (secondary shoots) to undergo early senescence, drying out and tissue
52 death. Ultimately, the significant loss in leaf area due to infection and death affects the plant’s
53 photosynthetic ability, leading to reductions in biomass and yield. There are no known rice
54 cultivars fully resistant to leaf sheath blight disease³⁻⁵. In the United States, planting high-
55 production rice varieties that are susceptible to SB, resulted in up to 50% yield losses³, in a
56 severe breakout season. As a result, about 72% of the rice planting area in the United States is
57 sprayed with fungicide that incurs unsustainable economic cost to the farmers⁶. The spraying
58 may also pose severe ecological constraints on animals and microbiome that co-inhabit and/or
59 frequently visit rice farms and fresh water bodies. Classified as a soilborne basidiomycete
60 fungus, *R. solani* is a particularly destructive plant pathogen⁷. In addition to its impact on rice, *R.*
61 *solani* infects other major crops such as soybean, barley, sorghum, tomato, and maize⁸. Although
62 *R. solani* rarely produces spores for a mode of mobility, sclerotia are formed which can survive
63 in the soil for up to two years⁹. Flooding in paddy fields, as common practice or natural
64 occurrence combined with a highly humid environment, the sclerotia spread and attach to the
65 plant, causing SB in the susceptible varieties.

66 Currently, disease management for SB relies heavily on fungicides. This approach may
67 not be the most sustainable due to potential resistance developing in *R. solani* populations, and
68 pollution of agricultural resources⁸. For example, in 2012, a strain of *R. solani* was isolated in the
69 state of Louisiana, USA which was resistant to the strobilurin class of fungicides used by rice
70 farmers^{10,11}. A more sustainable approach to disease management is to breed genetic resistance to
71 SB in commercial varieties. In this study, mRNA sequencing was utilized to compare gene
72 expression profiles of resistant and susceptible rice plants inoculated with *R. solani*. Using the
73 transcriptome profiles of two rice lines, susceptible Cocodrie (CCDR)¹², and resistant
74 MCR010277 (MCR)¹³ (Figure-1A), we examined the response to *R. solani* infection at multiple
75 time points including Day-0 (untreated control), and Day-1, 3, and 5 after inoculation.

76 The high-quality rice reference genome of *O. sativa japonica* cv Nipponbare (IRGSP
77 v1.0)¹⁴, bioinformatics analysis tools and sequences from public databases were used to profile
78 variety and time point-specific differential expression of all rice genes with special focus on
79 those overlapping and around the known major SB-resistance QTL on Chromosome 9¹⁵. We
80 aligned the transcriptome sequence reads to the reference rice genome to identify variety-
81 specific synonymous and non-synonymous SNPs, indels and additionally probed their putative
82 consequences on transcript structure, regulation and protein function *in-silico*. The select set of
83 SNP markers overlapping the differentially expressed candidate genes from the major SB-
84 resistance QTL region on chromosome-9 were confirmed by sequencing. The selected markers
85 were also genotyped in resistant and susceptible individuals derived from a biparental (CCDR
86 x MCR) doubled haploid (DH) population of 197 individuals.

87 We observed clear gene expression and phenotype differences in response to *R. solani*
88 infection in each of the two rice parental lines. Our findings also reveal plant defense response

89 genes that are differentially expressed in CCDR and MCR following inoculation and carry
90 genetic variation that alters the gene function. Based on the gene expression data, genotyping
91 and phenotyping, we present strong evidence supporting a rice *Wall Associated Kinase 91*
92 (OsWAK91) gene as a potential breeding target for developing SB resistant breeding lines.

93

94 **Results**

95 **Transcriptome analysis**

96 Twenty-four cDNA libraries were created, with 3 replicates from 4 time points for each rice line,
97 Day 0 (untreated control) and Day-1, 3, and 5 after *R. solani* inoculation. Strand-specific 150bp
98 paired-end sequencing of the cDNA libraries yielded 308,482,684 (CCDR) and 357,018,002
99 (MCR) reads. After filtering the low-quality raw reads, we obtained 299,687,236 and
100 346,498,184 (~98.5% of total) from the CCDR and MCR samples respectively, with an average
101 of 27.7 million reads per sample. The cDNA sequence reads were aligned to the *O. sativa*
102 *japonica* cv Nipponbare (IRGSP v1.0) reference genome. We found 16,480 (CCDR) and 17,793
103 (MCR) protein coding genes which show normalized expression from a minimum of one time
104 point. Expression of known genes from each line, at each time point, was compared in a pairwise
105 fashion against the untreated, Day-0 sample for differential gene expression analysis. For CCDR
106 there were 105, 367 and 377 transcripts from 79, 281 and 320 genes at the Day-0 to Day-1, Day-
107 0 to Day-3, and Day-0 to Day-5 comparisons, respectively. Whereas, in the MCR line, there
108 were 148 and 597 transcripts from 119 and 443 genes in the Day-0 to Day-3, and Day-0 to Day-5
109 comparisons, while there were no statistically significant differences in expression observed at
110 Day 0 to 1 (Figure-1B). In the Day-0 to Day-3 comparison, 17 common genes were up-regulated
111 in both the CCDR and MCR lines, and one common gene was down-regulated. Similarly, in the

112 Day 0 to 5 comparison, 63 common genes in the CCDR and MCR lines were up-regulated, while
113 8 genes were down-regulated (Figure 1B). In the susceptible CCDR line, there was a larger
114 number of differentially-expressed genes at Day-1, and increasing number on Day-3, unlike
115 the MCR line which has larger number of differentially expressed genes on Day-5 (Figure 1C).

116
117 Genes up-regulated in the CCDR line include transcription factors, and genes playing a role in
118 the biotic stress response (both at Day-0 to 1), and enrichment of hydrolase function in the Day-0
119 to Day-3 comparison. One gene (OS02G0129800) which is grass family-specific, was down-
120 regulated in both CCDR and MCR lines. This gene is known to be up-regulated in response to
121 *Burkholderia glumae*¹⁶ the causal agent of rice bacterial panicle blight disease, whereas it is
122 down-regulated in response to *Xanthomonas oryzae* the causal agent of rice bacterial leaf streak
123 disease¹⁷. The Day-0 to Day-5 comparison shows enrichment of genes associated with
124 catabolism of peptidoglycan, an important cell wall component, those associated with biotic
125 stimulus and defense responses, kinases and those with chitinase enzyme activity required for
126 degrading the fungal pathogen cell walls.

127
128 We found 1948 and 1914 genes from the CCDR and MCR lines, respectively, that undergo
129 alternative splicing of transcripts. Of these, 1890 genes were common to both lines. Of all the
130 genes showing alternative splicing under biotic stress, 57 and 20 unique genes were
131 differentially-expressed in the CCDR and MCR lines, respectively. Intron retention was
132 identified as the most prevalent form of transcript splicing event in response to *R. solani*
133 infection (Table-1). This is consistent with known reports from plants, including those under
134 abiotic stress¹⁸.

135
136 Four differentially-expressed candidate genes were identified in the vicinity of the major SB
137 resistance QTL on chromosome-9; a dormancy auxin-associated family protein
138 (OS09G0437500), an elicitor-inducible cytochrome P450 (OS09G0441400), a cysteine peptidase
139 (OS09G0442300) and a *Wall-Associated receptor kinase 91* (*WAK91*; OS09G0561600) with
140 contrasting expression profiles (Figure 1D). The *WAK91* transcript showed down-regulation up
141 to Day-3, followed by up-regulation at Day-5 in the MCR line compared to being constitutively
142 expressed in the CCDR line. OS09G0441400, the elicitor-responsive P450 gene, showed an
143 expression profile opposite that of *WAK91*. OS09G0437500, the dormancy auxin-associated
144 gene, exhibited a similar expression profile between the two lines, except at Day-1 in CCDR
145 when it was down-regulated. In contrast, OS09G0442300, the cysteine peptidase, was highly
146 expressed in MCR at Day-5.

147

148 **SNP marker discovery and variant prediction**

149 Using the transcriptome sequence reads, we identified 65,121 and 82,293 genetic variants,
150 including single nucleotide changes (SNPs in the form of transitions and transversions and
151 insertions and/or deletions (indels) in the susceptible CCDR and the resistant MCR lines,
152 respectively. A total of 32,401 and 49,573 SNPs were unique to CCDR and MCR, respectively,
153 and they shared 32,720 SNPs present at the same loci. Of the shared loci, 103 SNP sites
154 contained alleles different from the reference and between the two lines (Figure 1E). On average,
155 we observed that the MCR line carries ~2500 transitions (Ts), ~750 transversion (Tv) and ~1000
156 more indel SNP events compared to the CCDR line (Table-2). In general TvSNPs are have been
157 shown to have larger regulatory effects compared to the TsSNPs¹⁹, and rice is known to carry

158 more of the transition types of SNPs²⁰. The TsSNPs were greater in number in both lines, but the
159 transition to transversion ratio (Ts/Tv) in the MCR line was 2.8 compared to 2.68 in CCDR.
160 Within the TsSNPs, there were almost the same number of A ↔ G, and T ↔ C transitions in
161 CCDR compared to MCR, whereas there were more G → A and C → T transitions than A → G
162 and T → C, respectively. The transition type G→C was high in CCDR compared to MCR, where
163 the opposite type C→G is more abundant. There were significantly fewer TsSNP C→A in MCR
164 (Table 2).

165
166 The identified genetic variants map to 22,448 and 24,343 rice reference protein coding genes in
167 the CCDR and MCR lines respectively, constituting 55% and 66% of the genes in the published
168 reference genome¹⁴. Because we used the cDNA sequence reads for SNP calling and genetic
169 marker development, we used the Variant Effect Predictor (VEP) tool²¹ to predict the causal
170 effect of each SNP on the 5' and 3' UTRs, exons, intron and intron splicing boundaries of the
171 transcribed gene. If the SNP was in an exon, we also computed the synonymous (sSNP) and non-
172 synonymous (nsSNP) variations to predict the putative consequences on transcript structure,
173 regulation, splicing, and peptide structure and function (Table-3).

174
175 Based on these VEP results, and the differential gene expression results, we developed genetic
176 markers that co-segregate with the disease resistance phenotype in the population study. The set
177 of 135 SNP genetic markers²², including those identified herein, was short-listed by selecting for
178 chromosomal locations overlapping the major and minor SB QTL. Sheath blight QTLs were
179 previously identified on chromosomes 1, 2, 3, 4, 5, 6, 7, 8, 9, 11 and 12^{15,23-29}. In the major SB
180 QTL region, we found that the differentially expressed *OsWAK91* (OS09G0561600) gene carried

181 a non-synonymous SNP (T→C) aligned at position 22,318,449 bp on Chromosome-9 of the
182 Nipponbare rice reference genome (IRGSP v1.0; Figure 1F). The SB susceptible line CCDR
183 carried the same allele T as the reference, whereas the resistant MCR line carried the C allele.
184 The T→C transition results in the loss of a stop codon in MCR WAK91, resulting in a predicted
185 OsWAK91 peptide with an additional 68 amino acids. Sequencing the amplified region of the
186 WAK91 SNP marker confirmed for presence of the susceptible T allele in the parent CCDR, as
187 well as other US elite lines with japonica background, namely, CL53, Cypress, Blue Bonnet, and
188 CL111, whereas in the indica lines namely, IR29, IR64, Jasmine, TeQing, Pokkali, Nonabokra,
189 93-11, Kasalath and Nagina22 (N22) were confirmed to carry the resistant allele C (Figure-1G).

190

191 **Population study**

192 Evaluation of the 10 most-resistant individuals of the SB2 double haploid (DH) population
193 produced disease ratings between 4.7 to 6.0, while the resistant parent MCR showed a 3.5
194 rating in the same study. In contrast, the 10 most-susceptible individuals of the population
195 scored between 7.5 to 8.0, while the susceptible parent CCDR was a 7.5 rating (Figure-1A,
196 1H, 1I, Supplementary Data File 1). The selective genotyping of the parent lines MCR and
197 CCDR and 20 selected DH lines with 135 nsSNP markers and located in the major and minor
198 SB QTL regions on chromosomes 1, 2, 3, 4, 5, 6, 8, 9, 11 and 12. These included the non-
199 synonymous *OsWAK91* SNP (T/C) identified in this study. An ANOVA test was performed,
200 and the best 13 ranked nsSNP markers, based on the ANOVA analysis, F values, False
201 Discovery Rates, and adjusted R² values (Table 4), mapped to the bottom of chromosome-9,
202 within the major SB resistance QTL that was identified previously in 12 separate studies using
203 six *indica* lines (reviewed by Zuo et al. 2014³⁰). In the major SB QTL region on chromosome-

204 9, we found that the *OsWAK91* (OS09G0561600) gene carried a nsSNP (T/C) at position
205 22,318,449 bp on Chromosome-9 of the Nipponbare rice reference genome. Based on the
206 genotyping data we found that, except in the SB2-99 line, the *WAK91* SNP allele C
207 contributed by the MCR parent line always co-segregated with SB resistance phenotype,
208 whereas the allele T from the CCDR parent co-segregates with susceptible phenotype (Figure
209 1H, Supplementary Data File 1).

210

211 **Discussion**

212 In our quest of finding candidate genes in rice that can provide good measure of resistance to
213 *Rhizoctomia solani* and its application for breeding improved disease resistant rice, our study
214 integrated transcriptomics, identifying genetic variation and its causes on the gene function.
215 We identified the rice candidate gene *Wall-Associated Kinase 91* (*OsWAK91*) carrying a T/C
216 mutation at the stop codon. The *OsWAK91* SNP site at position 22,318,449 bp on
217 Chromosome-9 of the reference rice genome is part of stop codon (TAG) common to CCDR
218 and the reference japonica line Nipponbare. Due to the presence of allele C in the resistant line
219 MCR, the stop codon TAG is lost and replaced by the codon CAG in the *OsWAK91* open
220 reading frame. The codon CAG codes for glutamine, a polar uncharged side chain amino acid.
221 The stop codon loss in the MCR line also results in a *OsWAK91* peptide that is 68 amino
222 acids longer compared to that of CCDR and Nipponbare (Figure 1F). These results lead to two
223 hypotheses: (1) The *OsWAK91* SNP allele C is associated with SB resistance that is
224 evolutionarily conserved with *Oryza* ancestors, and (2) the longer *OsWAK91* ORF resulting
225 from the loss of the stop codon results in a peptide with gain in novel molecular function
226 which may play a role in leaf sheath blight resistance.

227

228 To test if the association of the *WAK91* SNP allele C with SB resistance is evolutionarily
229 conserved, we sequenced a PCR-amplified fragment overlapping the *WAK91* SNP loci from
230 13 additional *O. sativa* lines; five from subspecies *japonica* and eight from subspecies *indica*
231 (Figure-1G).

232 All US elite rice lines have the *japonica* background, and are known to be SB susceptible²².
233 We predicted, and confirmed that the US *japonica* lines CCDR, CL53, Cypress, Blue Bonnet,
234 CL111 carry the *OsWAK91* T allele associated with known susceptibility^{31,4,32-38}. In contrast,
235 the *indica* lines (IR29, IR64, Kasalath, Jasmin, TeQing, Nagina 22, Pokkali and Nonabokra)
236 were confirmed to carry the resistant C allele (Figure-1G), and exhibit some level of known
237 SB resistance phenotype^{35,39-41}. One exception is 93-11, from which the *indica* reference
238 genome was sequenced⁴², and which is known to be moderately susceptible^{39,43}. These results
239 are consistent with the genotype and phenotype observations of the 20 individuals of the SB2
240 DH population (with the exception of SB2-99), in which the *OsWAK91* SNP allele C
241 contributed by the MCR parent always co-segregated with SB resistance phenotype (Figure
242 1H, 1I, Supplementary Data File 1).

243

244 To gain insights on the origin of the *OsWAK91* SNP in the *Oryza* genus, we used synteny data
245 provided by the Gramene database⁴⁴ to query the *OsWAK91* SNP allele T/C by aligning to the
246 sequenced genomes of the wild species; *O. longistaminata*, *O. glaberrima*, *O. punctata*, *O.*
247 *meridionalis*, *O. barthii*, *O. glumaepatula*, *O. rufipogon*, *O. brachyantha* and the outgroup
248 *Leersia perrieri*. All of these species carry the *indica*-type allele C, and the majority are
249 known to bear the SB resistance phenotype (Figure-1G)^{3,33,45,46}. This suggests that the *indica*-

250 type resistant C allele is of ancestral origin. Conversely, the japonica-type susceptible T allele
251 is a more recent introduction in the *Oryza* clade, and is restricted to the *O. sativa subspecies*
252 *japonica* for its origin. We further confirmed its origin by mining the publicly-available
253 genetic variation data from the 3000 Rice Genome Project⁴⁷. The majority of the temperate
254 and tropical japonica and aromatic rice lines carry the japonica T allele, whereas the majority
255 of indica and AUS lines carry the C allele. Only a handful of the lines carry the heterozygous
256 T/C allele (Figure 1J). We did not have access to the SB resistance phenotype data for these
257 lines, however, based on our findings, we predict that the indica lines carrying the C allele
258 may bear some degree of SB resistance.

259

260 The second hypothesis suggests that the *OsWAK91* gene from the SB resistant MCR parent
261 encodes a peptide predicted to be 68 amino acid longer due to stop loss in the ORF, and this
262 peptide may have gained a function that plays a role in providing the SB resistance phenotype.
263 The *OsWAK91* gene is a member of the Wall-Associated receptor Kinase (WAK) gene family
264 (Figure 1K) and codes for a plasma membrane protein. The OsWAK91 protein contains a
265 wall-associated receptor kinase galacturonan-binding domain (WAK domain), followed by a
266 calcium-binding epidermal growth factor (EGF)-like domain in the N-terminal half of the
267 protein that is present on the extra-cellular side, in the middle is a single transmembrane
268 domain spanning the plasma membrane, followed by the cytosolic side containing a protein
269 kinase domain and additional phosphorylation sites for serine, threonine and tyrosine residues,
270 towards its C-terminal (Figure 1F). The WAK domain is known to be linked to the pectin
271 fraction of the plant cell wall⁴⁸⁻⁵⁰.

272

273 Members of the WAK gene family are known to express and function in response to biotic and
274 abiotic stress conditions. Expression of *Arabidopsis thaliana* *AtWAK1* is induced by a
275 pathogen response, by exogenous salicylate or by its analog 2,2-dichloroisonicotinic acid, and
276 requires the positive regulator NPR1/NIM1. The expression of complete *AtWAK1* or just its
277 kinase domain alone, can provide resistance to pathogen *Pseudomonas syringae*⁵¹. Maize
278 *ZmWAK-RLK* a plasma membrane protein, is known for conferring resistance to northern
279 corn leaf blight disease caused by the fungal pathogen *Setosphaeria turcica* (anamorph
280 *Exserohilum turcicum*; previously known as *Helminthosporium turcicum*), by restricting
281 pathogen entry into the host cells²⁵. Similarly wheat WAK gene *Stb6* confers fungal pathogen
282 resistance without a hypersensitive response⁵². Barley, *HvWAK1* is known to play a role in
283 root development; however, when compared to other cereals and *Arabidopsis* sequence
284 divergence in the extracellular domain further verifies the multifunctionality of WAK genes⁵³.
285 The WAKs are known to play role in plant cell expansion during seedling development, MAP
286 kinase signaling, and to bind to pectin polymers in the plant cell wall and have higher affinity
287 to bind smaller pectin fibers in response to pathogen attack⁴⁸.

288

289 In the reference rice genome, the WAK gene family members are distributed in 23 subfamilies
290 (Figure 1K) and are reported to play roles in development⁵⁴, abiotic responses^{55,56} and fungal
291 disease responses⁵⁷. Genes from five WAK subfamilies C, F, L, and Q did not show
292 expression in our study (Figure 1K). The gene *OsWAK1* (OS01G0136400) is known to
293 provide resistance to the rice blast disease pathogen *Magnaporthe grisea*, and its expression is
294 induced by salicylic acid and methyl jasmonate. *OsWAK1* phosphorylates itself and *OsRFP1*,
295 a putative transcription regulator which binds to the *OsWAK1* kinase domain. The rice

296 *OsWAK91* gene, also known as the *DEFECT IN EARLY EMBRYO SAC1 (OsDEES1)*, and
297 *OsWAK1* are known for playing role in rice sexual reproduction by regulating the
298 development of the female gametophyte (the embryo sac), and in anther dehiscence,
299 respectively^{54,58}. Overexpression of *OsWAK25* (OS03G0225700) is known to increase
300 susceptibility to *R. solani* and *Cochliobolus miyabeanus*, and resistance to *Xanthomonas*
301 *oryzae* pv. *oryzae* (*Xoo*) and *Magnaporthe oryzae*⁵⁹, whereas its loss of function compromises
302 Xa21-mediated resistance^{60,61}. *OsWAK11* has also been shown to regulates copper
303 detoxification^{55,56}.

304

305 Increased rice blast resistance and reduced fungal hyphae growth was observed when
306 *OsWAK91* was overexpressed, in contrast to the overexpression of *OsWAK112*
307 (OS10G0180800), which suppressed resistance. The genes *OsWAK14* (OS02G0632800),
308 *OsWAK91* and *OsWAK92* (OS09G0562600) are part of the heterotrimeric WAK protein
309 complex located on the plasma membrane in rice cells^{62,63} (Figure 1L). A mutant screening by
310 Delteil et al.⁵⁷ reported that the loss of function of *OsWAK14*, *OsWAK91* and *OsWAK92*
311 resulted in reduced basal resistance, though it did not affect the growth and fertility⁵⁷. .

312 Conversely, we concluded that both the CCDR and the Nipponbare rice reference lines carry
313 the susceptible T allele with shorter *OsWAK91* ORF, and are known for SB susceptibility. We
314 hypothesize that if the *OsWAK91* function is knocked out or when the longer C-terminal is
315 not present, the line should become increasingly pathogen susceptible. Therefore, we can
316 expect that the Nipponbare *OsWAK91* knock out line should be highly susceptible. The
317 results published by Delteil et al.⁵⁷ support this hypothesis, through testing with the pathogen
318 *Magnaporthe grisea*, causal agent of rice blight. In their study, Tos17-transposon knockout

319 mutant plants were generated in the japonica reference Nipponbare rice background
320 (susceptible). The wildtype Nipponbare and the Tos-17 mutant plants always showed disease
321 lesions in response to the pathogen, and, as expected, the *OsWAK91* mutant plant showed 2.5
322 fold increase in lesions.

323

324 The ability to detect and fight pathogenic microbes by the host plants is a complex process.
325 The pattern-recognition receptors (PRRs) localized on the surface of plant cells are well-
326 known gene family members of the Receptor-Like Kinase (RLK) proteins that assist in the
327 recognition of PAMP (Pathogen Associated Molecular Patterns) conserved motifs from
328 pathogens and DAMPs (Damage Associated Molecular Patterns) derived from the damage by
329 the microbes development^{64,65,63}. Also, in the case of *Magnaporthe grisea* blast disease
330 response in rice, the chitin elicitor binding protein, OsCEBiP and a chitin elicitor receptor
331 kinase, OsCERK, are known to form hetero-dimer complex. Upon chitin binding the CEBiP-
332 CERK receptor complex may in turn induce transphosphorylation followed by activating
333 downstream signaling⁶⁶ (Figure 1L). The majority of the OsCEBiP and OsCERK gene family
334 members show positive upregulation in the resistant MCR line (Supplementary file 1).
335 Similarly, the OsWAKs show early transcriptional regulation induced by chitin that is partially
336 under the control of the chitin receptor CEBiP⁵⁷. Query of the publicly available SNP datasets
337 ^{47,67,44} suggests that *OsWAK91* is known to carry ~400 unique SNP and indels present in the 5'
338 and 3' UTRs, introns, exons, and splice junctions. This includes the stop lost at position
339 22,318,449 bp identified by us, and indel at 9:22,315,774-22,315,791 bp leading to start lost
340 and 5'UTR variant, and stops gained at positions 22,316,748 bp, 22,317,724 bp, 22,318,006
341 bp on chromosome 9 (Supplementary Data File-2). These genetic differences did not show up

342 in our transcriptome-based SNP dataset and are potential genetic markers for further study.
343 Therefore, based upon the supporting evidence from the genotyping, phenotyping,
344 evolutionary ancestry of SB tolerance and the favorably segregating SNP in the SB2 DH
345 population, we propose a working model (Figure 1L) that the candidate gene *OsWAK91*
346 provides broad tolerance to leaf sheath blight similar to when infected with pathogens *R.*
347 *solani* and *Magnaporthe grisea*. The longer ORF of the *OsWAK91* in the indica lines, or the
348 MCR genotype with allele C, *OsWAK91* provides a complete kinase domain and additional
349 active sites for serine, threonine and tyrosine phosphorylation, which are missing in the
350 susceptible CCDD genotype, or japonica lines with shorter C-terminus (Figure 1F).
351 *OsWAK91* forms a heterotrimeric protein complex with *OsWAK14* and *OsWAK92* proteins,
352 and is an essential member of the WAK protein complex. The shorter kinase domain, and C-
353 terminus is expected to maintain proper embryo sac development and fertility in both the
354 japonica and indica lines. The *OsWAK91* in the japonica background with truncated C-
355 terminus kinase domain is not completely dysfunctional, since, it has been shown to be
356 involved in embryo sac development and the japonica line Nipponbare and others mentioned
357 in this work are all fertile. Whereas, the presence of the ancestral allele C contributes to the
358 stop lost in the candidate gene *OsWAK91*, the longer translated protein kinase domain and
359 additional phosphorylation sites in the C-terminus in indica lines are necessary for their role in
360 actively regulating the WAK protein complex function, and downstream signaling to affect the
361 disease tolerance response.

362

363

364

365 **Conclusions**

366 We determined and confirmed strong association of the *OsWAK91* SNP marker to the leaf sheath
367 blight resistance phenotype by extending the genotyping and phenotyping study to include 10
368 best resistant and 10 best susceptible individuals from the double haploid population derived
369 from the CCDR and MCR parents. We also mined the publicly available sequenced genomes
370 of reference and ancestral *Oryza* species and the 3000 rice genome project to confirm the
371 evolutionary source of the mutant SNP. The resistant allele and the trait appears ancestral,
372 whereas the susceptible allele is a more recent acquisition in the *O. sativa japonica* clade. In fact
373 all the US elite rice varieties have japonica genome integration and are known to carry the
374 susceptibility trait. A few of them tested by us confirm the presence of the susceptible SNP allele
375 same as the CCDR line. Our results and inferences, supported by sequencing, genotyping and
376 phenotyping experiment are well complemented by mutant/knockout screening by earlier
377 studies. *OsWAK91* knockout mutants make rice disease susceptible. Therefore *OsWAK91*
378 resistant allele and the genetic marker identified by us is of ancestral origin and is a candidate for
379 integration in the existing and new cultivated rice varieties for leaf sheath blight resistance,
380 which may help improve global rice production by rescuing the crop affected by the pathogen
381 *Rhizoctonia solani*

382

383 **Methods**

384

385 **Plant material and Inoculation**

386 The sheath blight resistant line MCR010277 and the susceptible variety Cocodrie (CCDR) were
387 planted in pots placed in a mist chamber in a greenhouse located on the Louisiana State

388 University (LSU) campus in Baton Rouge, LA. Plants were inoculated 50 days after germination
389 with a Potato Dextrose Agar PDA medium disc (0.8 cm diameter) containing *Rhizoctonia solani*
390 (strain LR172) mycelia. Discs were placed at the base of the stem and between the leaf blade and
391 leaf sheath on the main culm of each plant.

392 Temperatures inside the mist chamber ranged from a minimum of 27⁰C at night to a maximum
393 of 37⁰C in the day. Natural daylight was used, with day length of approximately 11 hours 30
394 minutes. Humidity was maintained 80-90% using a cool mist humidifier of 1.2 gallons capacity
395 (Vicks Inc.) that was programmed to function for a two-hour period every six hours. The
396 chamber frame was constructed with ¾ inch PVC pipe (Charlotte Pipe[®]) covered by extra light
397 plastic (0.31 mm) (Painter's Plastic – Poly America). The dimensions of the chamber were: 1.32
398 m wide by 2.70 m length, and 1.42 m in height, (Figure 4.2) for a total capacity of 48 pots per
399 chamber, each pot containing three plants.

400

401 Leaf samples approximately 2 cm in length were collected from the control untreated (day 0) and
402 from the inoculation sites 1, 3 and 5 days after treatment, and placed immediately in liquid
403 nitrogen until transfer to -80⁰C freezer for storage.

404

405 Leaf samples from additional 13 rice lines: CL153, Cypress, Blue Bonnet, CL111 (all japonica),
406 and 93-11, IR29, IR64, Kasalath, Jasmin, TeQing, Nagina 22 (N22), Pokkali and Nonabokra (all
407 indica) were collected for DNA extraction, and SNP marker genotyping by DNA amplification
408 and sequencing.

409

410

411

412 **RNA extraction and sequencing**

413 Frozen leaf samples from the CCDR and MCR lines collected at LSU were shipped to Oregon
414 State University on dry ice and stored at -80°C - until further processing. Poly(A)-enriched
415 mRNA libraries was prepared from total RNA extracted from the leaf samples using RNA Plant
416 Reagent[®] (Invitrogen Inc., USA), RNeasy kits (Qiagen Inc., USA) and RNase-free DNase (Life
417 Technologies Inc., USA). The concentration and quality of the poly(A)-enriched mRNA was
418 determined using a ND-100 spectrophotometer (Thermo Fisher Scientific Inc., USA) and
419 Bioanalyzer 2100 (Agilent Technologies Inc., USA), respectively. TruSeq RNA Sample
420 Preparation kits (Illumina Inc., USA) were used to construct sequencing libraries. An Illumina
421 HiSeq 3000 (Illumina Inc., USA) at the Center for Genome Research and Biocomputing, Oregon
422 State University (CGRB, OSU), was used to sequence the 150 bp Paired End cDNA libraries.

423

424 **Sequence quality control and read alignments**

425 The program Sickle v1.33 was used to filter all reads based on read quality. Reads under the
426 phred score 30, and read length under 150bp were rejected. FASTQ files containing quality-
427 trimmed and filtered reads were generated, yielding MCR and CCDR high-quality reads.
428 Sequence reads were further filtered by removing any reads showing alignment to the pathogen
429 *Rhizoctonia solani* AG1-IB isolate (GCA_000832345.1). Quality-filtered reads were aligned to
430 the reference *Oryza sativa* japonica cv Nipponbare (IRGSP v1.0) genome and annotation. Reads
431 from each biological replicate, were aligned with alignment software program STAR v.2.4.1a⁶⁸.
432 The resulting sequence alignment files were converted to binary alignment files (BAM), sorted
433 coordinately and indexed using SAMTools v.1.3⁶⁹.

434

435 **Differential gene expression (DGE)**

436 The program DESeq2 (v 1.22.2) was used to investigate gene expression levels of the two rice
437 lines over the time course⁷⁰. A Python package, HTSeq v.0.6.1p1⁷¹, was used to index the
438 alignment BAM files and generate raw counts, with a minimum quality-score cut-off of 10. The
439 DESeq2 R-package was the used to identify differentially expressed genes. First, the raw count
440 data was transformed using variance stabilizing transformation (vst) method of DESeq2 to filter
441 outliers. The DESeq2 functions namely, estimate size factors and estimate dispersions, were used
442 to normalize the aligned read counts. This was followed by the DESeq2 function to fit a negative
443 binomial general linear model and Wald test statistics to detect the significance scores (p-value).
444 The final set of differentially expressed transcripts were called at the FDR, Benjamin Hoffman,
445 corrected cutoff P-value of 0.05. Furthermore, the non-significant events identified by the
446 DESeq2 software were filtered out. Principle components (PCA) plots were used to visualize
447 cross-sample comparison in each DE pairwise comparison. MA-plots were generated to visualize
448 the different between signal intensity between samples in each comparison and dispersion plots
449 were generated to visualize the fitted normalized counts.

450

451 **SNP discovery and variant prediction**

452 Indexed BAM files were analyzed to identify genetic variations including single nucleotide
453 changes (SNPs; transitions and transversions) and insertions and/or deletions (indels).
454 Alignments from each variety were pooled using the mpileup function of SAMtools⁶⁹.
455 VarScan.v2.3.9 (release 80)⁷² was used to identify SNPs and indels with four different
456 minimum read coverage of categories of 2, 6, 8, and 20 and default minimum variant

457 frequency of 0.8 with a p-value of 0.005. A consensus set of SNPs were identified in all four
458 minimum read coverages. A genotype-specific SNP output data file in VCF format was
459 generated. The Variant Effect Predictor (VEP) workflow provided by the Gramene database²¹
460 was used to infer the putative consequences to the structure, splicing, and function of the gene
461 product (transcript and/or peptide) based on synonymous and/or nonsynonymous (ns) changes.

462

463 In order to develop genetic markers and build a hypothesis, the significant SNPs which
464 overlapped the major SB-resistance QTL on Chromosome-9 were queried to find association
465 to the observed differentially-expressed candidate genes. These same SNPs were also queried
466 against the 3000 Rice Genome data provided by the Rice SNP Seek project^{73,74} to profile
467 allelic nature and distribution in the diverse *O. sativa* lines. Some of the SNPs were also
468 queried on the Gramene database against the sequenced genomes of wild species of the *Oryza*
469 genus to infer presence or absence. Finally, peer-reviewed literature was mined to extract, and
470 confirm information on the known leaf sheath blight resistance phenotype for these rice lines.

471

472 **Scoring alternative splicing events**

473 SpliceGrapher, a Python-based tool⁷⁵ was used to score alternative splicing events. This tool
474 contains splice-site classifier files for each major reference genome, including *Oryza sativa*
475 *japonica*. These classifiers are used with the first program (sam_filter.py) in the SpliceGrapher
476 pipeline that intakes aligned experimental files and the provided classifier and outputs
477 alignment files filtered by splice site. Next, the filtered alignment files and the reference gene-
478 annotation file (in the gff file format), is used by the second program (predict_graphs.py) to
479 produce a gff file for each gene and its spliced forms. To compare isoforms detected in our

480 samples to the known reference isoforms, we utilized the program
481 (`gene_model_to_splicegraph.py`) to generate gffs for each known reference gene. The
482 SpliceGrapher pipeline provided statistics and graphical data files. The types of alternative
483 splicing events generated using SpliceGrapher are alternative 5' (5'-Alt) and 3' (3'-Alt)
484 events, exon skipping (ES) and intron retention (IR).

485

486 **Population study**

487 The RiceCAP SB2 mapping population was developed as a genetic resource to identify lines
488 containing molecular markers associated with SB resistance⁷⁶. The SB2 population consists of
489 197 doubled-haploid (DH) lines derived from a cross between the susceptible parent Cocodrie
490 (CCDR)¹² and the resistant parent MCR010277 (MCR)¹³. The population and the two parents
491 were evaluated for leaf sheath blight disease response on a scale of 0-9, where 0 = no disease
492 and 9 = dead plant. Plants were grown and phenotyped across three years at Crowley, LA and
493 Stuttgart, AR, USA²⁶. The top 10 SB resistant individuals SB2-03 (GSOR200003), SB2-109
494 (GSOR200109), SB2-134 (GSOR200134), SB2-158 (GSOR200158), SB2-161
495 (GSOR200161), SB2-174 (GSOR200174), SB2-206 (GSOR200206), SB2-225
496 (GSOR200225), SB2-259 (GSOR200206), and SB2-272 (GSOR200272) and the 10 most
497 susceptible individuals, SB2-13 (GSOR200013), SB2-48 (GSOR200048), SB2-88
498 (GSOR200088), SB2-99 (GSOR200099), SB2-125 (GSOR200125), SB2-144
499 (GSOR200144), SB2-203 (GSOR200203), SB2-255 (GSOR200255), SB2-276 (GSOR200276),
500 and SB2-314 (GSOR200314) from the SB2 DH population along with the two parents were
501 screened with 135 candidate nsSNP markers (Table 4, Supplementary Data File 3) identified
502 by Silva et al. 2012²², including the *OsWAK9I* SNP. These markers were located in QTL

503 regions on chromosomes 1, 2, 3, 4, 5, 6, 7, 8, 9, 11 and 12^{15,23-29}. An ANOVA analysis was
504 performed to calculate the F values, Bonferroni Hochberg False Discovery Rates (FDR), p
505 values and adjusted R² values.

506

507 **DNA extraction, amplification, sequencing and analysis**

508 Total DNA was extracted from seedling tissue using a DNeasy Plant Mini Kit (Qiagen Inc.,
509 USA). PCR primers were synthesized by Invitrogen Inc., USA for amplification of the region
510 of interest (Table 5). For PCR amplification, we used Ready PCR Mix, 2x (VWR) following
511 the recommended protocol on a C100 Thermocycler (BioRad). The thermocycler was set for 3
512 minutes at 95C for initial denaturation, followed by 40 cycles of 30 seconds at 95C for
513 denaturation, 30 seconds at 55C for annealing, 30 seconds at 72C for extension, and 5 minutes
514 at 72C for final extension. Amplified DNA fragments were separated on a 1% agarose gel in
515 1x TAE buffer. These fragments were extracted from the gel using QIAGEN Gel Extraction
516 and QIAGEN PCR product cleanup kits. Isolated PCR products were Sanger sequenced at the
517 CGRB, OSU. The amplified sequences from the PCR products for each experimental line,
518 including MCR and CCDR, were aligned to the reference Nipponbare genome (IRGSP v1.0)
519 using ClustalW for confirmation of SNPs and to identify and confirm the alleles.

520

521 **Data mining and functional annotation**

522 The reference genome sequences, gene function and gene family annotations were
523 downloaded in bulk and/or queried online at Gramene database (www.gramene.org)⁴⁴. The
524 3000 rice genome project data was mined at Rice SNP-Seek Database ([http://snp-](http://snp-seek.irri.org/)
525 [seek.irri.org/](http://snp-seek.irri.org/))^{73,74}. Gramene database was queried to pull out all rice genes (IRGSP v1.0)

526 annotated to carry the WAK domain and/or were part of WAK gene family. The protein
527 sequences of the canonical (longest) isoform) were downloaded in fasta format to run the
528 CLUSTALW alignment with default parameters. The gene family tree and the differential
529 gene expression data for the respective rice WAK gene family members were uploaded on the
530 iTOL web portal⁷⁷ to generate the graphics (Figure 1K).

531

532 **Data Availability:** The raw RNA sequence reads were submitted to the EMBL-EBI
533 ArrayExpress (accession number E-MTAB-6402). The raw sequences from the amplified
534 fragments from various rice lines were submitted to the European Nucleotide Archive (study
535 accession PRJEB28811) under the assigned accession numbers ERS2758541-51.

536

537 **References from Zotero (nature style):**

- 538 1. Lau, W. The Challenges of Rice Sustainability: the Increasing World's Population and
539 Climate Change. *The Challenges of Rice Sustainability: the Increasing World's Population*
540 *and Climate Change* (2017).
- 541 2. Lee, F. N. Rice Sheath Blight: A Major Rice Disease. *Plant Dis.* **67**, 829 (1983).
- 542 3. Prasad, B. & Eizenga, G. C. Rice Sheath Blight Disease Resistance Identified in *Oryza* spp.
543 Accessions. *Plant Dis.* **92**, 1503–1509 (2008).
- 544 4. Pan, X. B. *et al.* Major Gene, Nonallelic Sheath Blight Resistance from the Rice Cultivars
545 Jasmine 85 and Teqing. *Crop Sci.* **39**, 338–346 (1999).
- 546 5. Singh, R., Sunder, S. & Kumar, P. Sheath blight of rice: current status and perspectives.
547 *Indian Phytopathol.* **69**, 12 (2016).
- 548 6. Nalley, L., Tsiboe, F., Durand-Morat, A., Shew, A. & Thoma, G. Economic and
549 Environmental Impact of Rice Blast Pathogen (*Magnaporthe oryzae*) Alleviation in the
550 United States. *PLoS ONE* **11**, (2016).
- 551 7. Zheng, A. *et al.* The evolution and pathogenic mechanisms of the rice sheath blight
552 pathogen. *Nat. Commun.* **4**, 1424 (2013).
- 553 8. Zhang, C. Q., Liu, Y. H., Ma, X. Y., Feng, Z. & Ma, Z. H. Characterization of sensitivity of
554 *Rhizoctonia solani*, causing rice sheath blight, to mepronil and boscalid. *Crop Prot.* (2009).
- 555 9. Webster, editor.), Robert K., Gunnell, editor.), Pamela S., APS Press, publisher. &
556 American Phytopathological Society, issuing body. *Compendium of rice diseases*. (St. Paul,
557 Minn.□: APS Press, 1992).
- 558 10. OLAYA, G. Detection of resistance to QoI fungicides in *Rhizoctonia solani* isolates from
559 rice. in (2012).

- 560 11. Fungicide-resistant *Rhizoctonia solani* found in Louisiana. *Farm Progress* (2012). Available
561 at: [https://www.farmprogress.com/rice/fungicide-resistant-rhizoctonia-solani-found-](https://www.farmprogress.com/rice/fungicide-resistant-rhizoctonia-solani-found-louisiana)
562 [louisiana](https://www.farmprogress.com/rice/fungicide-resistant-rhizoctonia-solani-found-louisiana). (Accessed: 26th March 2019)
- 563 12. Linscombe, S. D. *et al.* Registration of ‘Cocodrie’ Rice. *Crop Sci.* **40**, 294–294 (2000).
- 564 13. Rush, M., Nandakumar, R., Zhang, S. & Groth, D. Status of our research on developing high
565 levels of partial resistance to sheath blight in lines with high yield potential and good
566 agronomic characteristics. in (2006).
- 567 14. Kawahara, Y. *et al.* Improvement of the *Oryza sativa* Nipponbare reference genome using
568 next generation sequence and optical map data. *Rice N. Y. N* **6**, 4 (2013).
- 569 15. Li, Z., Pinson, S. R., Marchetti, M. A., Stansel, J. W. & Park, W. D. Characterization of
570 quantitative trait loci (QTLs) in cultivated rice contributing to field resistance to sheath
571 blight (*Rhizoctonia solani*). *TAG Theor. Appl. Genet. Theor. Angew. Genet.* **91**, 382–388
572 (1995).
- 573 16. Magbanua, Z. V. *et al.* Transcriptomic dissection of the rice – *Burkholderia glumae*
574 interaction. *BMC Genomics* **15**, 755 (2014).
- 575 17. Wilkins, K. E., Booher, N. J., Wang, L. & Bogdanove, A. J. TAL effectors and activation of
576 predicted host targets distinguish Asian from African strains of the rice pathogen
577 *Xanthomonas oryzae* pv. *oryzicola* while strict conservation suggests universal importance
578 of five TAL effectors. *Front. Plant Sci.* **6**, (2015).
- 579 18. Filichkin, S. A. *et al.* Environmental Stresses Modulate Abundance and Timing of
580 Alternatively Spliced Circadian Transcripts in *Arabidopsis*. *Mol. Plant* **8**, 207–227 (2015).
- 581 19. Guo, C. *et al.* Transversions have larger regulatory effects than transitions. *BMC Genomics*
582 **18**, (2017).

- 583 20. Li, G. *et al.* Genome-Wide Sequencing of 41 Rice (*Oryza sativa* L.) Mutated Lines Reveals
584 Diverse Mutations Induced by Fast-Neutron Irradiation. *Mol. Plant* **9**, 1078–1081 (2016).
- 585 21. Naithani, S., Geniza, M. & Jaiswal, P. Variant Effect Prediction Analysis Using Resources
586 Available at Gramene Database. *Methods Mol. Biol. Clifton NJ* **1533**, 279–297 (2017).
- 587 22. Silva, J. *et al.* Identification of candidate genes in rice for resistance to sheath blight disease
588 by whole genome sequencing. *Theor. Appl. Genet.* **124**, 63–74 (2012).
- 589 23. Sato, H. *et al.* Mapping QTLs for Sheath Blight Resistance in the Rice Line WSS2. *Breed.*
590 *Sci.* **54**, 265–271 (2004).
- 591 24. Pinson, S. R. M., Capdevielle, F. M. & Oard, J. H. Confirming QTLs and Finding Additional
592 Loci Conditioning Sheath Blight Resistance in Rice Using Recombinant Inbred Lines
593 Mention of a trademark or proprietary product does not constitute a guarantee or warranty of
594 the product by the USDA and does not imply its approval to the exclusion of other products
595 that also can be suitable. *Crop Sci.* **45**, 503–510 (2005).
- 596 25. Zeng, Y., Ji, Z., Ma, L., Li, X. & Yang, C. Advances in Mapping Loci Conferring Resistance
597 to Rice Sheath Blight and Mining *Rhizoctonia solani* Resistant Resources. *Rice Sci.* **18**, 56–
598 66 (2011).
- 599 26. Nelson, J. C. *et al.* Sheath-blight resistance QTLs in japonica rice germplasm. *Euphytica*
600 **184**, 23–34 (2012).
- 601 27. Jia, Y. *et al.* Registration of Four Rice Germplasm Lines with Improved Resistance to Sheath
602 Blight and Blast Diseases. *J. Plant Regist.* **6**, 95 (2012).
- 603 28. Taguchi-Shiobara, F. *et al.* Mapping and validation of QTLs for rice sheath blight resistance.
604 *Breed. Sci.* **63**, 301–308 (2013).

- 605 29. Costanzo, S., Jackson, A. K. & Brooks, S. A. High-resolution mapping of Rsn1, a locus
606 controlling sensitivity of rice to a necrosis-inducing phytotoxin from *Rhizoctonia solani*
607 AG1-IA. *Theor. Appl. Genet.* **123**, 33–41 (2011).
- 608 30. Zuo, S. *et al.* Fine-mapping of qSB-9TQ, a gene conferring major quantitative resistance to
609 rice sheath blight. *Mol. Breed.* **34**, 2191–2203 (2014).
- 610 31. Masajo, T. M. Varietal Resistance in Rice to Sheath Blight Caused by *Thanatephorus*
611 *Cucumeris* (Frank) Donk (*Rhizoctonia Solani* Kuhn). 169
- 612 32. Wamishe, Y., Cartwright, R. & Lee, F. Management of Rice Diseases. in *Compendium of*
613 *Rice Diseases and Pests* 16
- 614 33. Eizenga, G. C., Lee, F. N. & Rutger, J. N. Screening *Oryza* Species Plants for Rice Sheath
615 Blight Resistance. *Plant Dis.* **86**, 808–812 (2002).
- 616 34. Groth, D. E. & Bond, J. A. Effects of Cultivars and Fungicides on Rice Sheath Blight, Yield,
617 and Quality. *Plant Dis.* **91**, 1647–1650 (2007).
- 618 35. Jia, Y. *et al.* Rapid Determination of Rice Cultivar Responses to the Sheath Blight Pathogen
619 *Rhizoctonia solani* Using a Micro-Chamber Screening Method. *Plant Dis.* **91**, 485–489
620 (2007).
- 621 36. Srinivasachary, Willocquet, L. & Savary, S. Resistance to rice sheath blight (*Rhizoctonia*
622 *solani* Kühn) [(teleomorph: *Thanatephorus cucumeris* (A.B. Frank) Donk.)] disease: current
623 status and perspectives. *Euphytica* **178**, 1–22 (2011).
- 624 37. XiaoXing, L., Ping, L. & AiPing, Z. In vitro selection of sheath blight resistance germplasm
625 in rice. *Afr. J. Microbiol. Res.* **7**, 4422–4429 (2013).
- 626 38. The International Rice Genebank Collection Information System. Available at:
627 <http://irgcis.irri.org:81/grc/SearchData.htm>. (Accessed: 29th March 2019)

- 628 39. Zuo, S. *et al.* Prospect of the QTL-qSB-9Tq utilized in molecular breeding program of
629 japonica rice against sheath blight. *J. Genet. Genomics* **35**, 499–505 (2008).
- 630 40. Hossain, M. K. *et al.* Identification and validation of sheath blight resistance in rice (*Oryza*
631 *sativa* L.) cultivars against *Rhizoctonia solani*. *Can. J. Plant Pathol.* **36**, 482–490 (2014).
- 632 41. Chaudhary, B. Evaluating Sheath Blight Resistance in Rice Using Detached Tiller and Field
633 Screening Method. *J. Nepal Agric. Res. Counc.* **1**, (2015).
- 634 42. Yu, J. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**,
635 79–92 (2002).
- 636 43. Xia, Y. *et al.* Transcriptome analysis reveals the host selection fitness mechanisms of the
637 *Rhizoctonia solani* AG11A pathogen. *Sci. Rep.* **7**, 10120 (2017).
- 638 44. Tello-Ruiz, M. K. *et al.* Gramene 2018: unifying comparative genomics and pathway
639 resources for plant research. *Nucleic Acids Res.* **46**, D1181–D1189 (2018).
- 640 45. Ram, T. *et al.* Identification of donors for sheath blight resistance in wild species of rice.
641 *Indian J. Genet. Plant Breed. India* (2008).
- 642 46. Djedatin, G. *et al.* Evaluation of African Cultivated Rice *Oryza glaberrima* for Resistance to
643 Bacterial Blight. *Plant Dis.* **95**, 441–447 (2011).
- 644 47. Li, J.-Y., Wang, J. & Zeigler, R. S. The 3,000 rice genomes project: new opportunities and
645 challenges for future rice research. *GigaScience* **3**, 8 (2014).
- 646 48. Kohorn, B. D. & Kohorn, S. L. The cell wall-associated kinases, WAKs, as pectin receptors.
647 *Front. Plant Sci.* **3**, (2012).
- 648 49. Anderson, C. M. *et al.* WAKs: cell wall-associated kinases linking the cytoplasm to the
649 extracellular matrix. *Plant Mol. Biol.* **47**, 197–206 (2001).

- 650 50. Wagner, T. A. & Kohorn, B. D. Wall-Associated Kinases Are Expressed throughout Plant
651 Development and Are Required for Cell Expansion. *Plant Cell* **13**, 303–318 (2001).
- 652 51. He, P. *et al.* Overexpression of Pti5 in Tomato Potentiates Pathogen-Induced Defense Gene
653 Expression and Enhances Disease Resistance to *Pseudomonas syringae* pv. tomato. *Mol.*
654 *Plant. Microbe Interact.* **14**, 1453–1457 (2001).
- 655 52. Saintenac, C. *et al.* Wheat receptor-kinase-like protein Stb6 controls gene-for-gene resistance
656 to fungal pathogen *Zymoseptoria tritici*. *Nat. Genet.* **50**, 368 (2018).
- 657 53. Kaur, R., Singh, K. & Singh, J. A root-specific wall-associated kinase gene, HvWAK1,
658 regulates root growth and is highly divergent in barley and other cereals. *Funct. Integr.*
659 *Genomics* **13**, 167–177 (2013).
- 660 54. Kanneganti, V. & Gupta, A. K. Wall associated kinases from plants — an overview. *Physiol.*
661 *Mol. Biol. Plants Int. J. Funct. Plant Biol.* **14**, 109–118 (2008).
- 662 55. Hu, W. *et al.* Cloning and characterization of the *Oryza sativa* wall-associated kinase gene
663 OsWAK11 and its transcriptional response to abiotic stresses. *Plant Soil* **384**, 335–346
664 (2014).
- 665 56. Xia, Y. *et al.* OsWAK11, a rice wall-associated kinase, regulates Cu detoxification by
666 alteration the immobilization of Cu in cell walls. *Environ. Exp. Bot.* **150**, 99–105 (2018).
- 667 57. Delteil, A. *et al.* Several wall-associated kinases participate positively and negatively in basal
668 defense against rice blast fungus. *BMC Plant Biol.* **16**, 17 (2016).
- 669 58. Wang, N. *et al.* The Rice Wall-Associated Receptor-Like Kinase Gene OsDEES1 Plays a
670 Role in Female Gametophyte Development. *Plant Physiol.* **160**, 696–707 (2012).
- 671 59. Harkenrider, M. *et al.* Overexpression of Rice Wall-Associated Kinase 25 (OsWAK25)
672 Alters Resistance to Bacterial and Fungal Pathogens. *PLOS ONE* **11**, e0147310 (2016).

- 673 60. Seo, Y.-S. *et al.* Towards Establishment of a Rice Stress Response Interactome. *PLOS Genet.*
674 7, e1002020 (2011).
- 675 61. Ke, Y., Deng, H. & Wang, S. Advances in understanding broad-spectrum resistance to
676 pathogens in rice. *Plant J.* **90**, 738–748 (2017).
- 677 62. Li, L., Yu, Y., Zhou, Z. & Zhou, J.-M. Plant pattern-recognition receptors controlling innate
678 immunity. *Sci. China Life Sci.* **59**, 878–888 (2016).
- 679 63. Cayrol, B., Delteil, A., Gobbato, E., Kroj, T. & Morel, J.-B. Three wall-associated kinases
680 required for rice basal immunity form protein complexes in the plasma membrane. *Plant*
681 *Signal. Behav.* **11**, e1149676 (2016).
- 682 64. Boller, T. & Felix, G. A renaissance of elicitors: perception of microbe-associated molecular
683 patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* **60**, 379–
684 406 (2009).
- 685 65. Macho, A. P. & Zipfel, C. Plant PRRs and the activation of innate immune signaling. *Mol.*
686 *Cell* **54**, 263–272 (2014).
- 687 66. Hayafune, M. *et al.* Chitin-induced activation of immune signaling by the rice receptor
688 CEBiP relies on a unique sandwich-type dimerization. *Proc. Natl. Acad. Sci.* **111**, E404–
689 E413 (2014).
- 690 67. Duitama, J. *et al.* Whole Genome Sequencing of Elite Rice Cultivars as a Comprehensive
691 Information Resource for Marker Assisted Selection. *PLOS ONE* **10**, e0124617 (2015).
- 692 68. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–
693 21 (2013).
- 694 69. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–
695 2079 (2009).

- 696 70. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
697 RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- 698 71. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-
699 throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
- 700 72. Koboldt, D. C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in
701 cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
- 702 73. Mansueto, L. *et al.* Rice SNP-seek database update: new SNPs, indels, and queries. *Nucleic*
703 *Acids Res.* **45**, D1075–D1081 (2017).
- 704 74. Wang, W. *et al.* Genomic variation in 3,010 diverse accessions of Asian cultivated rice.
705 *Nature* **557**, 43 (2018).
- 706 75. Rogers, M. F., Thomas, J., Reddy, A. S. & Ben-Hur, A. SpliceGrapher: detecting patterns of
707 alternative splicing from RNA-Seq data in the context of gene models and EST data.
708 *Genome Biol.* **13**, R4 (2012).
- 709 76. Chu, Q. *et al.* Registration of a C/M doubled haploid mapping population of rice. **46**, 1417
710 (2006).
- 711 77. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and
712 annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
- 713
- 714

715 **Acknowledgments**

716 We would like to thank Center for Genome Research and Biocomputing (CGRB) core facility
717 staff, Mark Dasenko for Illumina cluster generation and sequencing, Matthew Peterson, Chris
718 Sullivan and Justin Elser for computational support and Dr. Laurel Cooper and Justin Preece
719 for English language editing. Funding for the project was supported by NSF awards (IOS-
720 1127112 and IOS- 1340112) and the [Louisiana Rice Research Board](#).

721

722 **Author Contributions**

723 PJ and JO conceived and designed the experiments. YSG, MG, NAB and AM performed the
724 experiments. NAB, MG, AM, YSG, PJ, JO analyzed the data. Everyone wrote the paper:

725

726 **Competing Interest**

727 The Authors declare no competing interests.

728

729

730 **Figure Legend:**

731

732 **Figure-1:**

733

734 (A) The rice leaf sheath blight disease response shown by the resistant parent MCR010277

735 (MCR), the susceptible parent Cocodrie (CCDR) and the SB2-174 individual of the SB2

736 double haploid population. The disease response was tested by infecting the plants with the

737 fungal pathogen *Rhizoctonia solani* (strain LR172).

738 (B) The number of differentially expressed genes between the two parent lines when

739 compared to the Day-0 vs Day-1, Day-3 and Day-5 after infection.

740 (C) Number of up and down regulated genes shared between the time points within each rice

741 line.

742 (D) The differential expression of four candidate genes overlapping the major leaf sheath

743 resistance QTL on Chromosome 9 of rice. Rows are each of the four candidate genes and

744 columns are (L-R) first three, MCR Day-0 vs Day-1, Day-0 vs Day-3, Day-0 vs Day-5 and last

745 three, CCDR Day-0 vs Day-1, Day-0 vs Day-3, and Day-0 vs Day-5

746 (E) Number of unique and shared SNPs and indels identified in the CCDR and the MCR lines.

747 (F) The OsWAK91 ORF showing the regions of the coded domains (top) and zoom-in view of

748 the CDS and the ORF from the reference Nipponbare, and parents CCDR and MCR rice lines

749 (bottom). The presence of allele C in the leaf sheath blight resistant MCR line results in stop

750 codon loss and 82 amino acid longer WAK91 peptide which carries a longer protein kinase

751 domain(cyan colored) and additional predicted sites for serine (S), threonine (T) and tyrosine

752 (Y) phosphorylation sites (green colored).

753 (G) Genotyping and Phenotyping of the top ten resistant and ten susceptible individuals of the
754 SB2 double haploid population. The population and the two parents were evaluated for sheath
755 blight response phenotype on a scale of 0-9, where 0 = no disease and 9 = dead plant. The
756 OsWAK91 SNP alleles T/C were scored for presence/absence.

757 (H) OsWAK91 SNP allele T/C in 13 rice lines was confirmed by sequencing the SNP.
758 Synteny-based sequence data was mined at the publicly available Gramene database for the
759 reference lines indica 93-11, japonica Nipponbare, the wild species of *Oryza* genus and its
760 outgroup *Leersia Perrieri*. The leaf sheath blight disease response phenotypes were mined
761 from the previously published literature.

762 (I) The leaf sheath blight response phenotype of the individuals SB2-276 (susceptible) and
763 SB2-03 (resistant) from the DH population.

764 (J) Occurrence of the OsWAK91 SNP allele T/C The genotyping for in the genomes of the
765 3000 rice lines and their grouping

766 (K) Phylogenetic tree of the rice Wall associated Kinase gene family members distributed in
767 23 Subfamilies. Differential gene expression data from the MSC and CCDR rice lines were
768 plotted in the track order (outside to inside) MCR Day-0 Vs Day-1, Day-3, Day-5 and CCDR
769 Day-0 Vs Day-1, Day-3 and Day-5.

770 (K) A model of the host-pathogen interaction showing response to fungal chitin binding by the
771 chitin binding receptor complex, and potential interactions and functions by the heterotrimeric
772 WAK receptor protein complex. The WAK complex is functional in both the MCR and the
773 CCDR rice lines, however the MCR OsWAK91 C-terminus carrying the longer protein kinase
774 domain and the additional serine, threonine and tyrosine phosphorylation sites is expected to

775 play a role in successfully initiating the downstream signaling response to provide leaf sheath

776 blight resistance.

777

778

779 **Tables:**

780

781

782 Table-1

783 Summary of the number of genes and events undergoing alternative splicing of transcripts in the
784 susceptible CCDR and resistant MCR rice lines at different time points.

785

786

Rice line	MCR							
Time points	Day-0		Day-1		Day-3		Day-5	
Splicing type	Genes	Events	Genes	Events	Genes	Events	Genes	Events
Alt 5' UTR	335	392	339	396	335	392	335	392
Alt 3' UTR	553	652	553	652	551	650	551	650
Exon Skipping (ES)	189	204	191	207	185	200	185	200
Intron retention (IR)	1,357	1,651	1,357	1,651	1,350	1,644	1,351	1645
Rice line	CCDR							
Time points	Day-0		Day-1		Day-3		Day-5	
Splicing type	Genes	Events	Genes	Events	Genes	Events	Genes	Events
Alt 5' UTR	337	394	336	393	335	392	336	393
Alt 3' UTR	557	656	553	652	551	650	553	652
Exon Skipping (ES)	186	201	189	204	187	202	185	200
Intron retention (IR)	1,371	1,666	1,369	1,664	1,355	1,649	1,357	1,651

787

788

789

790 **Table 2**

791 Counts of SNPs and indels identified in the susceptible CCDR and the resistant MCR rice lines,
792 by aligning the RNA-Seq sequence reads against the reference *O. sativa japonica* cv Nipponbare
793 reference genome (IRGSP v1.0).

794

SNP	Substitution Type	CCDR	MCR
A → G	Transition	8,703	11,102
T → C	Transition	8,896	11,228
G → A	Transition	8,735	11,420
C → T	Transition	8,878	11,622
G → C	Transversion	2,817	3,523
C → G	Transversion	2,730	3,559
T → G	Transversion	3,274	4,108
G → T	Transversion	3,367	4,158
A → C	Transversion	3,316	3,964
C → A	Transversion	3,339	4,175
T → A	Transversion	3,734	4,428
A → T	Transversion	3,685	4,397
Total SNPs		61,474	77,684
Total Indels		3,647	4,609
Total variants		65,121	82,293

795

796

797 **Table-3**

798 Summary of SNP consequences on gene function and structure based on its overlapping position
799 in the genes identified in the reference *O. sativa japonica* cv Nipponbare reference genome
800 (IRGSP v1.0).

801

SNP Consequence	CCDR	MCR
splice donor variant	3,207	3,457
splice acceptor variant	1,952	2,167
stop gained	256	325
stop lost	96	124
start lost	26	47
missense variant	12,951	18,211
splice region variant	8,527	9,494
synonymous variant	12,849	17,534
stop retained variant	29	33
coding sequence variant	10	8
5 prime UTR variant	4,222	6,031
3 prime UTR variant	17,605	20,730
non coding transcript exon variant	2,389	2,694
intron variant	19,605	23,266
non coding transcript variant	2,854	3,311
upstream gene variant	99,220	124,576
downstream gene variant	110,185	138,011
intergenic variant	887	1,103

802

803

804 **Table: 4**

805 Results of the ANOVA analysis. Includes Marker rank, MSU (v7) and IRGSP v1.0 rice gene
 806 IDs, F value, Bonferroni Hochberg False Discovery Rate (FDR), p values and R-squared values
 807 of 135 nsSNP makers genotyped in the 10 most resistant and 10 most susceptible lines of the
 808 RiceCAP project's SB2 double haploid mapping population.
 809

Rank	MSU gene ID	IRGSP gene ID	F value	Raw p value	p value	FDR	R ² value
1	LOC_Os09g32860	N/A	149.15400	0.00000	0.00000	0.00000	0.89232
2	LOC_Os09g33710	Os09g0511900	61.04000	0.00000	0.00004	0.00000	0.77227
3	LOC_Os09g34180	Os09g0517600	61.04000	0.00000	0.00004	0.00000	0.77227
4	LOC_Os09g36900	Os09g0540600	61.04000	0.00000	0.00004	0.00000	0.77227
5	LOC_Os09g37230	Os09g0544300	61.04000	0.00000	0.00004	0.00000	0.77227
6	LOC_Os09g37240	Os09g0544400	61.04000	0.00000	0.00004	0.00000	0.77227
7	LOC_Os09g37590	N/A	61.04000	0.00000	0.00004	0.00000	0.77227
8	LOC_Os09g37800	Os09g0550600	61.04000	0.00000	0.00004	0.00000	0.77227
9	LOC_Os09g37880	Os09g0551400	61.04000	0.00000	0.00004	0.00000	0.77227
10	LOC_Os09g38700	Os09g0559900	61.04000	0.00000	0.00004	0.00000	0.77227
11	LOC_Os09g38710	Os09g0560000	61.04000	0.00000	0.00004	0.00000	0.77227
12	LOC_Os09g38850	Os09g0561600	61.04000	0.00000	0.00004	0.00000	0.77227
13	LOC_Os09g38970	Os09g0563250	61.04000	0.00000	0.00004	0.00000	0.77227
14	LOC_Os12g10330	Os12g0204600	41.63300	0.00000	0.00052	0.00004	0.69815
15	LOC_Os12g10410	Os12g0205500	41.63300	0.00000	0.00052	0.00004	0.69815
16	LOC_Os12g13100	Os12g0233100	41.63300	0.00000	0.00052	0.00004	0.69815
17	LOC_Os12g15460	Os12g0256900	41.63300	0.00000	0.00052	0.00004	0.69815
18	LOC_Os09g39620	Os09g0569800	37.14500	0.00001	0.00106	0.00008	0.67359
19	LOC_Os12g09710	Os12g0198900	33.11300	0.00002	0.00211	0.00016	0.56581
20	LOC_Os12g10180	N/A	33.11300	0.00002	0.00211	0.00016	0.56581
21	LOC_Os09g32020	Os09g0493500	23.45700	0.00013	0.01449	0.00097	0.48368
22	LOC_Os12g06980	Os12g0167700	23.45700	0.00013	0.01449	0.00097	0.48359
23	LOC_Os12g09000	Os12g0192500	16.85600	0.00066	0.07303	0.00468	0.44991
24	LOC_Os12g07950	Os12g0179800	14.72200	0.00121	0.13166	0.00807	0.35200
25	LOC_Os06g13040	Os06g0237400	9.77800	0.00583	0.62916	0.03699	0.33220
26	LOC_Os06g15170	Os06g0262800	8.95400	0.00781	0.83613	0.04726	0.28421
27	LOC_Os12g03554	Os12g0129550	7.14700	0.01550	0.98350	0.08949	0.26383
28	LOC_Os12g04660	Os12g0140700	7.14700	0.01550	0.98350	0.08949	0.26383
29	LOC_Os02g34490	Os02g0549900	6.28300	0.02201	0.98350	0.10351	0.25873
30	LOC_Os02g34850	Os02g0554000	6.28300	0.02201	0.98350	0.10351	0.25873
31	LOC_Os02g35210	Os02g0558400	6.28300	0.02201	0.98350	0.10351	0.25873

32	LOC_Os12g07800	Os12g0177800	6.45100	0.02053	0.98350	0.10351	0.25598
33	LOC_Os12g06740	Os12g0164300	6.45100	0.02053	0.98350	0.10351	0.24260
34	LOC_Os08g19694	Os08g0293300	6.19300	0.02284	0.98350	0.10361	0.24120
35	LOC_Os08g20020	Os08g0296900	5.72100	0.02788	0.98350	0.11423	0.24120
36	LOC_Os08g30850	Os08g0399300	5.72100	0.02788	0.98350	0.11423	0.18705
37	LOC_Os08g30910	Os08g0400200	5.76500	0.02736	0.98350	0.11423	0.18705
38	LOC_Os06g19110	Os06g0295000	3.85200	0.06535	0.98350	0.20748	0.17626
39	LOC_Os06g22020	Os06g0325900	3.85200	0.06535	0.98350	0.20748	0.17626
40	LOC_Os06g22460	Os06g0330100	3.85200	0.06535	0.98350	0.20748	0.17626
41	LOC_Os06g23530	Os06g0343100	3.85200	0.06535	0.98350	0.20748	0.17626
42	LOC_Os06g28124	Os06g0475400	3.85200	0.06535	0.98350	0.20748	0.17626
43	LOC_Os06g28670	Os06g0481400	3.85200	0.06535	0.98350	0.20748	0.17626
44	LOC_Os06g29700	Os06g0492800	3.85200	0.06535	0.98350	0.20748	0.17626
45	LOC_Os06g29844	Os06g0494400	4.14200	0.05684	0.98350	0.20748	0.17626
46	LOC_Os03g43684	N/A	4.14200	0.05684	0.98350	0.20748	0.16570
47	LOC_Os06g31070	Os06g0507200	3.57500	0.07486	0.98350	0.23189	0.12005
48	LOC_Os06g32350	N/A	2.20400	0.15499	0.98350	0.33529	0.12005
49	LOC_Os04g10460	Os04g0183500	2.20400	0.15499	0.98350	0.33529	0.11751
50	LOC_Os04g11640	Os04g0193200	2.20400	0.15499	0.98350	0.33529	0.11751
51	LOC_Os04g11970	Os04g0196200	2.20400	0.15499	0.98350	0.33529	0.11751
52	LOC_Os04g15650	Os04g0227000	2.39700	0.13899	0.98350	0.33529	0.11751
53	LOC_Os04g20680	Os04g0275100	2.39700	0.13899	0.98350	0.33529	0.11751
54	LOC_Os04g21890	N/A	2.39700	0.13899	0.98350	0.33529	0.11751
55	LOC_Os04g23620	N/A	2.39700	0.13899	0.98350	0.33529	0.11751
56	LOC_Os04g23890	Os04g0304200	2.39700	0.13899	0.98350	0.33529	0.11751
57	LOC_Os09g16540	Os09g0334800	2.39700	0.13899	0.98350	0.33529	0.10953
58	LOC_Os09g17600	Os09g0344800	2.39700	0.13899	0.98350	0.33529	0.10953
59	LOC_Os09g17630	Os09g0345300	2.39700	0.13899	0.98350	0.33529	0.10953
60	LOC_Os03g30130	Os03g0415200	2.45600	0.13451	0.98350	0.33529	0.10907
61	LOC_Os03g37720	Os03g0573500	2.45600	0.13451	0.98350	0.33529	0.10907
62	LOC_Os03g39150	Os03g0588400	2.16600	0.15840	0.98350	0.33529	0.10907
63	LOC_Os03g40250	Os03g0599400	2.16600	0.15840	0.98350	0.33529	0.10907
64	LOC_Os08g10560	Os08g0206500	2.21400	0.15407	0.98350	0.33529	0.10739
65	LOC_Os08g12800	Os08g0224500	2.21400	0.15407	0.98350	0.33529	0.10739
66	LOC_Os08g13870	Os08g0236400	2.21400	0.15407	0.98350	0.33529	0.10739
67	LOC_Os03g63110	Os03g0848400	1.47500	0.24022	0.98350	0.46294	0.07666
68	LOC_Os02g42412	Os02g0636400	1.47300	0.24058	0.98350	0.46294	0.07575
69	LOC_Os02g44730	Os02g0667500	1.47300	0.24058	0.98350	0.46294	0.07564
70	LOC_Os02g45160	Os02g0673100	1.47300	0.24058	0.98350	0.46294	0.07564

71	LOC_Os02g45980	Os02g0684900	1.47300	0.24058	0.98350	0.46294	0.07564
72	LOC_Os02g48210	Os02g0712700	1.49400	0.23730	0.98350	0.46294	0.07564
73	LOC_Os08g42930	Os08g0542200	1.20300	0.28720	0.98350	0.54439	0.06264
74	LOC_Os11g13650	N/A	1.04300	0.32068	0.98350	0.59891	0.06264
75	LOC_Os02g58540	Os02g0832150	0.92900	0.34792	0.98350	0.63123	0.05477
76	LOC_Os01g13300	Os01g0234100	0.92900	0.34792	0.98350	0.63123	0.04907
77	LOC_Os04g56250	Os04g0657600	0.89800	0.35586	0.98350	0.63654	0.04907
78	LOC_Os08g35310	Os08g0454500	0.85100	0.36855	0.98350	0.64118	0.04752
79	LOC_Os02g39590	Os02g0608801	0.85100	0.36855	0.98350	0.64118	0.04513
80	LOC_Os02g51900	Os02g0755500	0.75300	0.39685	0.98350	0.68108	0.04513
81	LOC_Os02g52060	Os02g0757400	0.61500	0.44303	0.98350	0.70947	0.04513
82	LOC_Os02g02650	Os02g0118700	0.53600	0.47334	0.98350	0.70947	0.04017
83	LOC_Os02g43460	Os02g0650800	0.53600	0.47334	0.98350	0.70947	0.03305
84	LOC_Os02g53970	Os02g0780200	0.53600	0.47334	0.98350	0.70947	0.02894
85	LOC_Os02g54330	N/A	0.53600	0.47334	0.98350	0.70947	0.02894
86	LOC_Os02g54500	Os02g0786000	0.53300	0.47484	0.98350	0.70947	0.02894
87	LOC_Os02g55180	Os02g0795000	0.53300	0.47484	0.98350	0.70947	0.02894
88	LOC_Os04g57670	Os04g0672700	0.53300	0.47484	0.98350	0.70947	0.02875
89	LOC_Os04g58720	Os04g0683800	0.53300	0.47484	0.98350	0.70947	0.02875
90	LOC_Os04g58820	Os04g0685000	0.53300	0.47484	0.98350	0.70947	0.02875
91	LOC_Os04g58910	Os04g0685900	0.53300	0.47484	0.98350	0.70947	0.02875
92	LOC_Os04g59060	Os04g0687300	0.50800	0.48531	0.98350	0.71668	0.02875
93	LOC_Os04g59540	Os04g0691900	0.43300	0.51884	0.98350	0.74688	0.02875
94	LOC_Os05g50660	Os05g0583500	0.42400	0.52340	0.98350	0.74688	0.02875
95	LOC_Os09g26300	Os09g0432900	0.42400	0.52340	0.98350	0.74688	0.02743
96	LOC_Os04g05030	Os04g0136900	0.34300	0.56525	0.98350	0.74778	0.02349
97	LOC_Os05g37040	Os05g0442100	0.34300	0.56525	0.98350	0.74778	0.02299
98	LOC_Os05g39760	N/A	0.35600	0.55839	0.98350	0.74778	0.02299
99	LOC_Os08g36320	Os08g0465800	0.36200	0.55499	0.98350	0.74778	0.02063
100	LOC_Os08g36760	Os08g0471800	0.37900	0.54578	0.98350	0.74778	0.02063
101	LOC_Os09g27570	Os09g0448100	0.37900	0.54578	0.98350	0.74778	0.02063
102	LOC_Os06g44820	Os06g0658600	0.37900	0.54578	0.98350	0.74778	0.01971
103	LOC_Os02g49986	Os02g0732600	0.13700	0.71588	0.98350	0.85611	0.01937
104	LOC_Os02g09820	Os02g0191500	0.13700	0.71588	0.98350	0.85611	0.01871
105	LOC_Os02g10120	Os02g0194700	0.13700	0.71588	0.98350	0.85611	0.01871
106	LOC_Os02g56380	Os02g0807900	0.12700	0.72610	0.98350	0.85611	0.00789
107	LOC_Os02g56480	Os02g0809100	0.14300	0.70964	0.98350	0.85611	0.00789
108	LOC_Os02g57960	Os02g0826000	0.14300	0.70964	0.98350	0.85611	0.00789
109	LOC_Os06g35850	Os06g0551800	0.14300	0.70964	0.98350	0.85611	0.00789

110	LOC_Os06g37500	Os06g0572300	0.14300	0.70964	0.98350	0.85611	0.00789
111	LOC_Os01g52330	Os01g0721300	0.11200	0.74152	0.98350	0.85611	0.00754
112	LOC_Os01g52880	Os01g0729400	0.11200	0.74152	0.98350	0.85611	0.00754
113	LOC_Os01g53420	Os01g0735900	0.11200	0.74152	0.98350	0.85611	0.00754
114	LOC_Os09g25620	Os09g0424300	0.14300	0.70964	0.98350	0.85611	0.00731
115	LOC_Os09g25890	Os09g0427800	0.14300	0.70964	0.98350	0.85611	0.00731
116	LOC_Os02g10900	Os02g0203500	0.13900	0.71698	0.98350	0.85611	0.00699
117	LOC_Os05g40790	Os05g0486300	0.13200	0.72010	0.98350	0.85611	0.00620
118	LOC_Os05g41130	Os05g0490300	0.13200	0.72010	0.98350	0.85611	0.00620
119	LOC_Os05g41290	Os05g0492200	0.13200	0.72010	0.98350	0.85611	0.00620
120	LOC_Os11g19700	Os11g0302500	0.04200	0.83928	0.98350	0.89571	0.00263
121	LOC_Os11g24060	Os11g0427700	0.04200	0.83928	0.98350	0.89571	0.00263
122	LOC_Os11g24180	Os11g0430000	0.04200	0.83928	0.98350	0.89571	0.00263
123	LOC_Os11g24770	N/A	0.04200	0.83928	0.98350	0.89571	0.00263
124	LOC_Os11g28950	Os11g0479100	0.04700	0.83001	0.98350	0.89571	0.00263
125	LOC_Os03g53220	Os03g0743800	0.04700	0.83001	0.98350	0.89571	0.00235
126	LOC_Os03g56400	Os03g0775400	0.04700	0.83001	0.98350	0.89571	0.00235
127	LOC_Os03g57160	Os03g0785300	0.04700	0.83001	0.98350	0.89571	0.00235
128	LOC_Os03g58390	Os03g0798200	0.04700	0.83001	0.98350	0.89571	0.00235
129	LOC_Os04g55760	Os04g0651500	0.03100	0.86218	0.98350	0.91248	0.00172
130	LOC_Os02g11820	Os02g0208900	0.00000	0.98350	0.98350	0.98350	0.00007
131	LOC_Os01g54350	Os01g0747400	0.00000	0.98350	0.98350	0.98350	0.00002
132	LOC_Os01g54515	Os01g0748950	0.00000	0.98350	0.98350	0.98350	0.00002
133	LOC_Os01g55050	Os01g0754500	0.00000	0.98350	0.98350	0.98350	0.00002
134	LOC_Os01g56040	Os01g0765900	0.00000	0.98350	0.98350	0.98350	0.00002
135	LOC_Os01g57230	Os01g0780600	0.00000	0.98350	0.98350	0.98350	0.00002
136	LOC_Os01g57900	Os01g0788900	0.00100	0.97305	0.98350	0.98350	0.00002

810

811

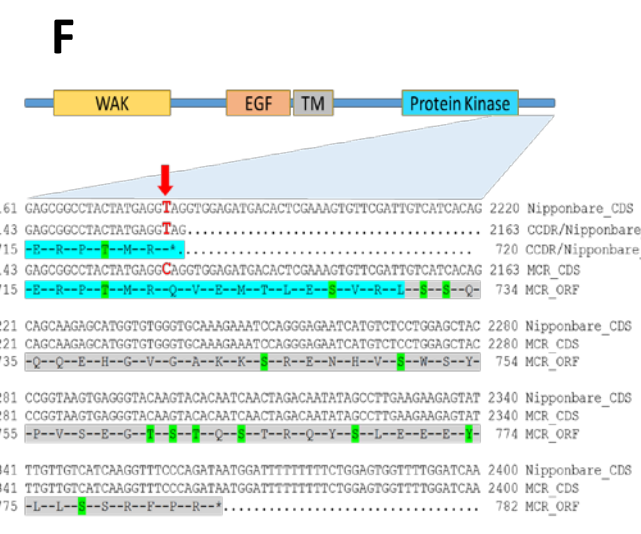
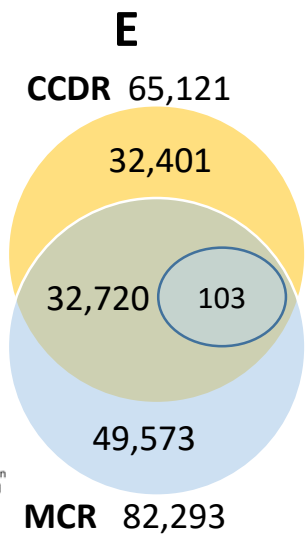
812 **Table 5** OsWAK SNP Primer Sets and PCR conditions. Primer name with FS and RS suffix
813 were used for amplifying the SNP region and Sanger sequencing. The primer names with
814 suffix FG were used for direct SNP-specific PCR-based genotyping.
815

Primer	Sequence	Length (nt)	Temperature Degree C	%GC
For PCR and Sanger sequencing				
Forward-FS	CAGGTCAAGGAAGAGGGAGG	20	59	60
Reverse-RS	AGGAGACATGATTCTCCCTGG	20	58	50
For PCR amplification-based genotyping				
Forward-FG1	GAACACTTTCGAGTGTCATCTC CACCAA	28	62	46
Forward-FG2	ACACTTTCGAGTGTCATCTCCA CC CG	26	64	54

816

817

818

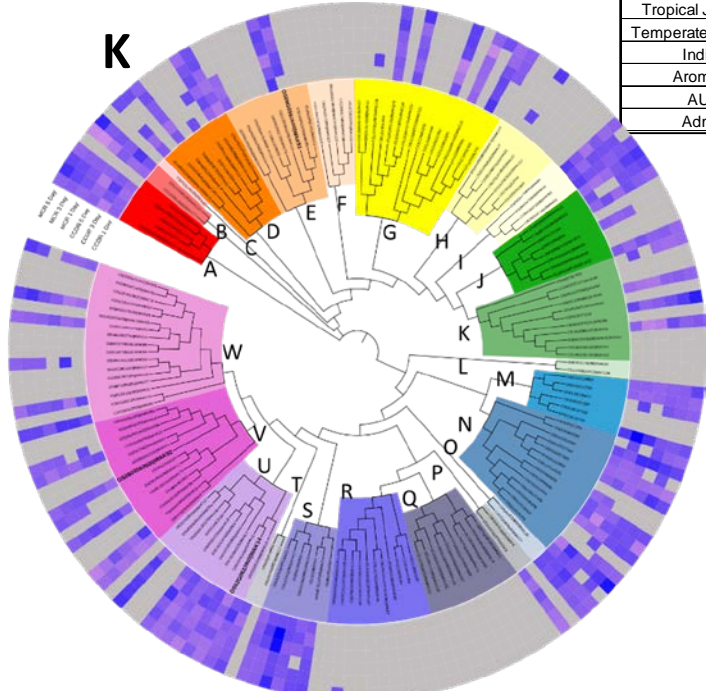
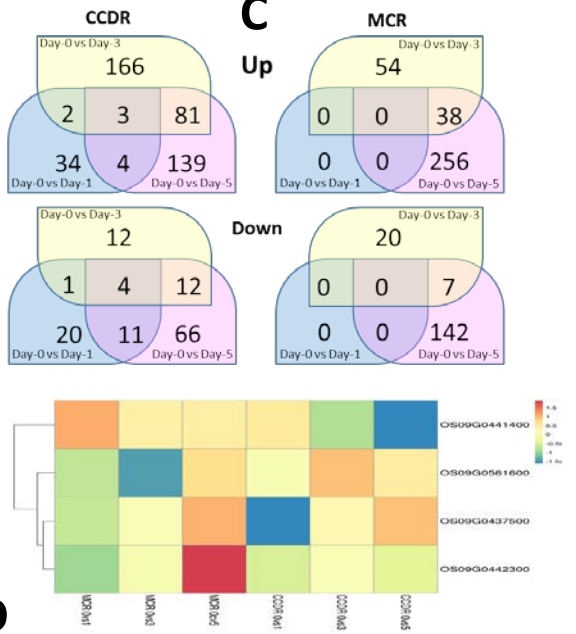
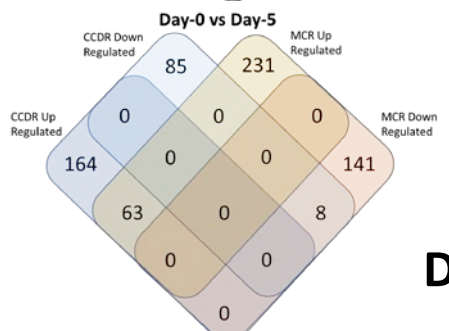
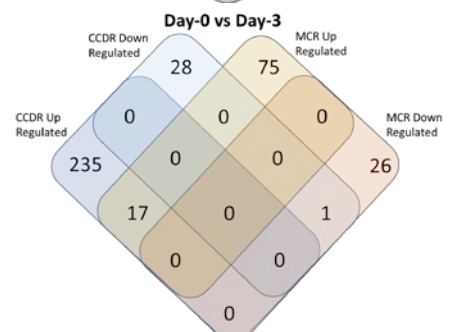
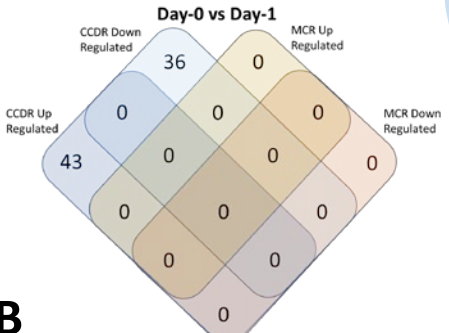


G

Genetic Diversity		
Rice line	Allele	Phenotype
Nipponbare (japonica reference)	T	Susceptible
CL153	T	Susceptible
Cypress	T	Susceptible
Blue Bonnet	T	Susceptible
CL111	T	Susceptible
93-11 (indica reference)	C	Moderate Susceptible
Nonabokra	C	Resistant
Pokkali	C	Resistant
Jasmine	C	Resistant
TeQing	C	Resistant
IR64	C	Resistant
IR29	C	Resistant
Nagina 22 (N22)	C	Resistant
Kasalath (AUS reference)	C	Resistant
<i>O. nivara</i>	C	Resistant
<i>O. rufipogon</i>	C	Resistant
<i>O. barthii</i>	C	Resistant
<i>O. brachyantha</i>	C	N/A
<i>O. glaberrima</i>	C	N/A
<i>O. glumaepatula</i>	C	Resistant
<i>O. longistaminata</i>	C	Resistant
<i>O. meridionalis</i>	C	Resistant
<i>O. punctata</i>	C	N/A
<i>Leersia perrieri</i>	C	N/A

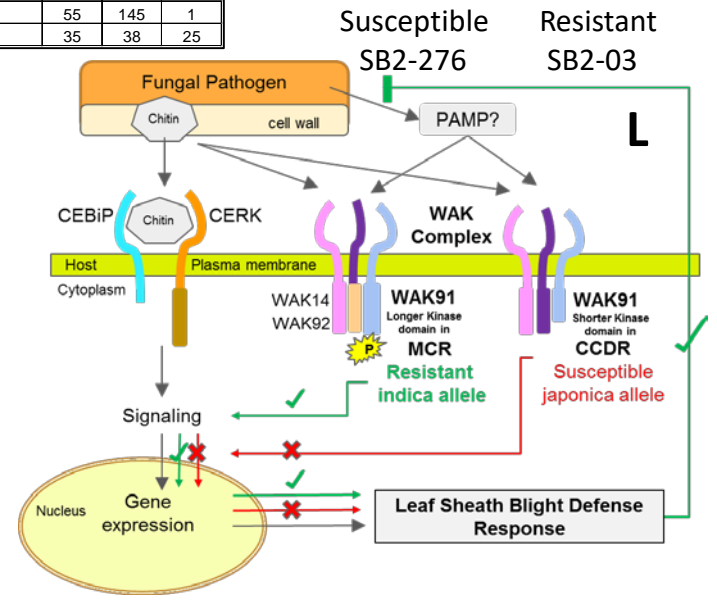
H

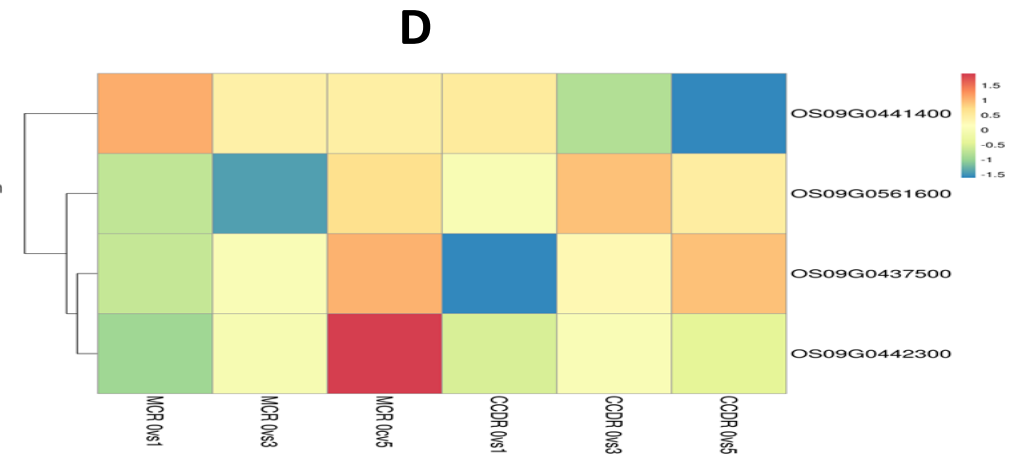
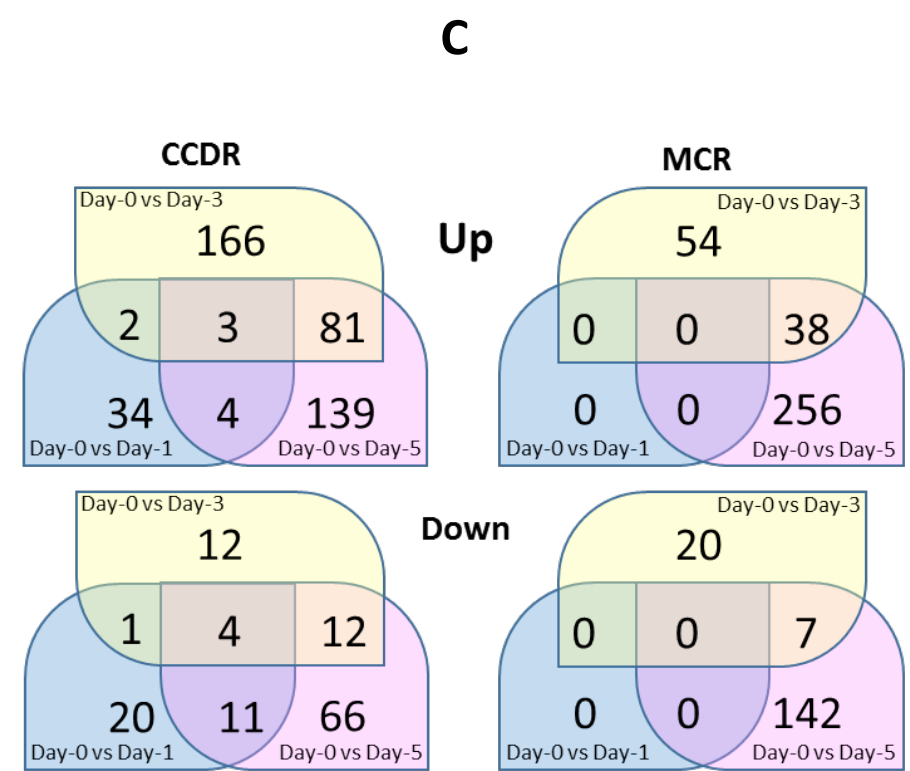
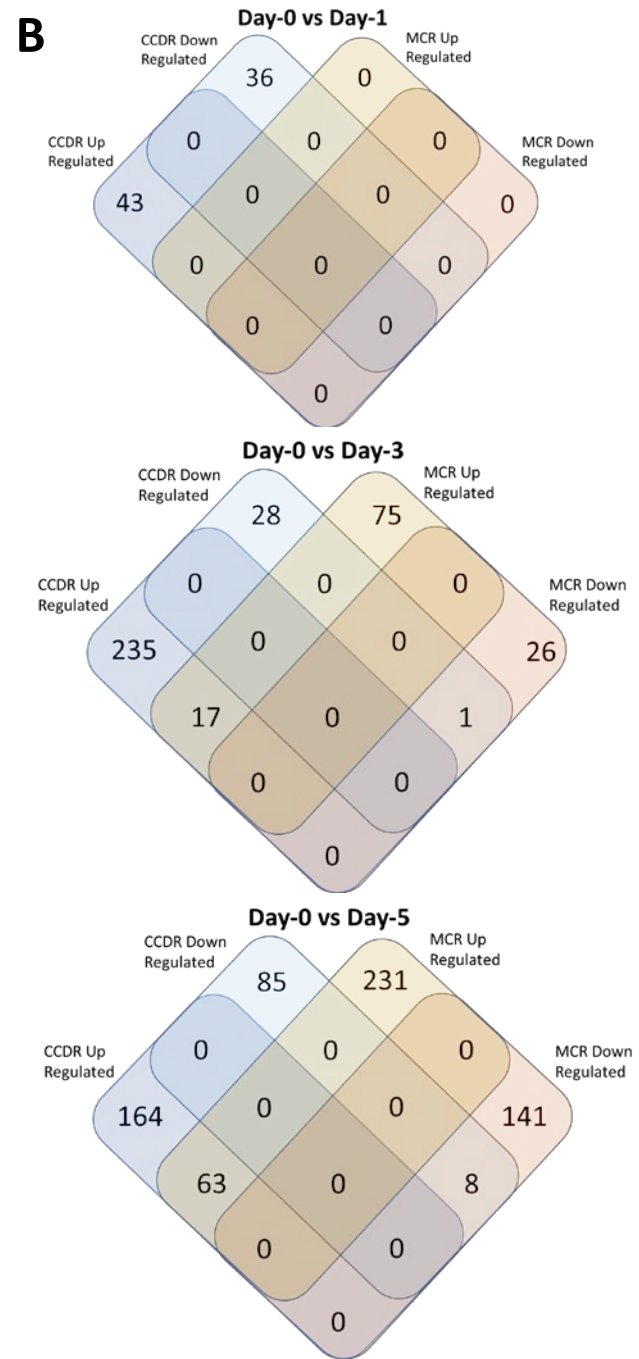
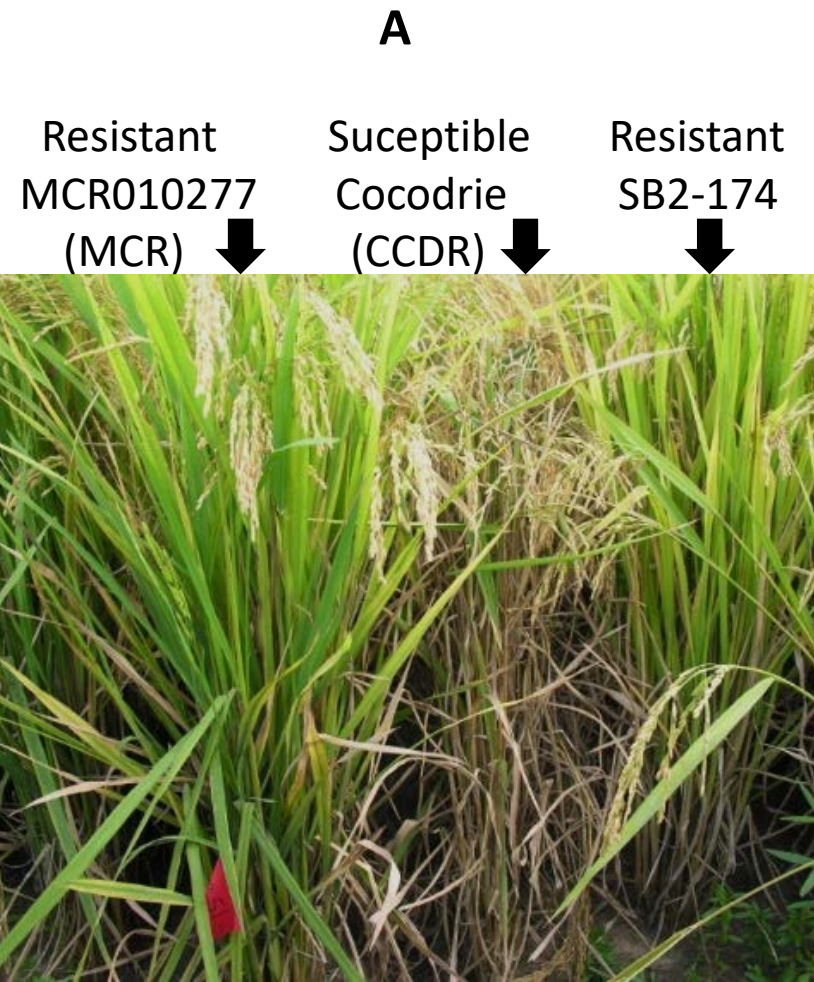
Population			
Individual	Allele	Phenotype	Score
CCDR (Parent)	T	Susceptible	7.5
SB2-99	C	Susceptible	7.5
SB2-13	T	Susceptible	7.5
SB2-48	T	Susceptible	8
SB2-88	T	Susceptible	8
SB2-126	T	Susceptible	8
SB2-144	T	Susceptible	8
SB2-203	T	Susceptible	7.5
SB2-255	T	Susceptible	8
SB2-276	T	Susceptible	8
SB2-314	T	Susceptible	7.5
MCR (Parent)	C	Resistant	3.5
SB2-03	C	Resistant	4.7
SB2-109	C	Resistant	5.5
SB2-134	C	Resistant	5.5
SB2-158	C	Resistant	5
SB2-161	C	Resistant	5.7
SB2-174	C	Resistant	4.5
SB2-259	C	Resistant	5.7
SB2-206	C	Resistant	6
SB2-225	C	Resistant	6
SB2-272	C	Resistant	6

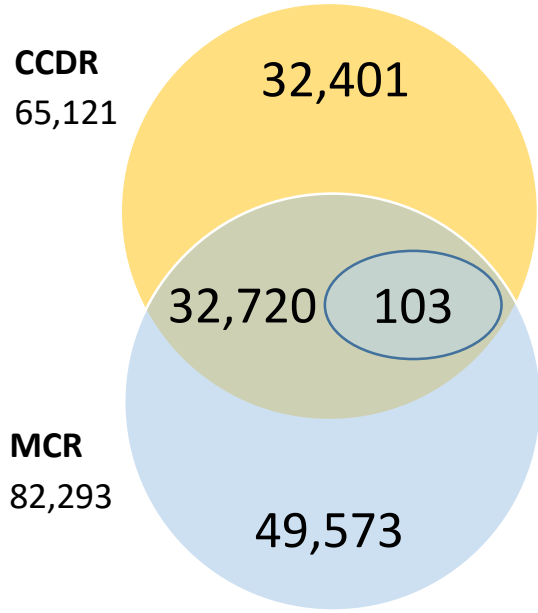
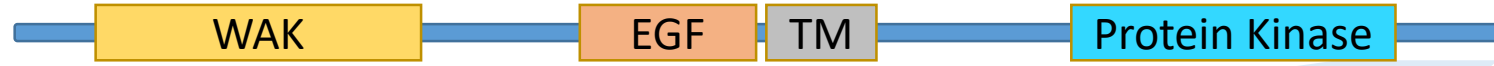


J

3K Rice genome project			
Type	Allele		
	T	C	T/C
Tropical Japonica	471	5	0
Temperate Japonica	360	0	3
Indica	18	1742	22
Aromatic	68	4	1
AUS	55	145	1
Admix	35	38	25





E**F**

2161 GAGCGGCCTACTATGAGG**T**AGGTGGAGATGACACTCGAAAGTGTTTCGATTGTCATCACAG 2220 Nipponbare_CDS
 2143 GAGCGGCCTACTATGAGG**T**AG..... 2163 CCDR/Nipponbare_CDS
 715 **-E--R--P--T--M--R--***..... 720 CCDR/Nipponbare_ORF

2143 GAGCGGCCTACTATGAGG**C**AGGTGGAGATGACACTCGAAAGTGTTTCGATTGTCATCACAG 2163 MCR_CDS
 715 **-E--R--P--T--M--R--Q--V--E--M--T--L--E--S--V--R--L--S--S--Q--** 734 MCR_ORF

2221 CAGCAAGAGCATGGTGTGGGTGCAAAGAAATCCAGGGAGAATCATGTCTCCTGGAGCTAC 2280 Nipponbare_CDS
 2221 CAGCAAGAGCATGGTGTGGGTGCAAAGAAATCCAGGGAGAATCATGTCTCCTGGAGCTAC 2280 MCR_CDS
 735 **-Q--Q--E--H--G--V--G--A--K--K--S--R--E--N--H--V--S--W--S--Y--** 754 MCR_ORF

2281 CCGGTAAGTGAGGGTACAAGTACACAATCAACTAGACAATATAGCCTTGAAGAAGAGTAT 2340 Nipponbare_CDS
 2281 CCGGTAAGTGAGGGTACAAGTACACAATCAACTAGACAATATAGCCTTGAAGAAGAGTAT 2340 MCR_CDS
 755 **-P--V--S--E--G--T--S--T--Q--S--T--R--Q--Y--S--L--E--E--E--Y--** 774 MCR_ORF

2341 TTGTTGTCATCAAGGTTTCCAGATAATGGATTTTTTTTTTCTGGAGTGGTTTTGGATCAA 2400 Nipponbare_CDS
 2341 TTGTTGTCATCAAGGTTTCCAGATAATGGATTTTTTTTTTCTGGAGTGGTTTTGGATCAA 2400 MCR_CDS
 775 **-L--L--S--S--R--F--P--R--***..... 782 MCR_ORF

G

Genetic Diversity		
Rice line	Allele	Phenotype
Nipponbare (japonica reference)	T	Susceptible
CL153	T	Susceptible
Cypress	T	Susceptible
Blue Bonnet	T	Susceptible
CL111	T	Susceptible
93-11 (indica reference)	C	Moderate Susceptible
Nonabokra	C	Resistant
Pokkali	C	Resistant
Jasmine	C	Resistant
TeQing	C	Resistant
IR64	C	Resistant
IR29	C	Resistant
Nagina 22 (N22)	C	Resistant
Kasalath (AUS reference)	C	Resistant
<i>O. nivara</i>	C	Resistant
<i>O. rufipogon</i>	C	Resistant
<i>O. barthii</i>	C	Resistant
<i>O. brachyantha</i>	C	N/A
<i>O. glaberrima</i>	C	N/A
<i>O. glumaepatula</i>	C	Resistant
<i>O. longistaminata</i>	C	Resistant
<i>O. meridionalis</i>	C	Resistant
<i>O. punctata</i>	C	N/A
<i>Leersia perrieri</i>	C	N/A

H

Population			
Individual	Allele	Phenotype	Score
CCDR (Parent)	T	Susceptible	7.5
SB2-99	C	Susceptible	7.5
SB2-13	T	Susceptible	7.5
SB2-48	T	Susceptible	8
SB2-88	T	Susceptible	8
SB2-126	T	Susceptible	8
SB2-144	T	Susceptible	8
SB2-203	T	Susceptible	7.5
SB2-255	T	Susceptible	8
SB2-276	T	Susceptible	8
SB2-314	T	Susceptible	7.5
MCR (Parent)	C	Resistant	3.5
SB2-03	C	Resistant	4.7
SB2-109	C	Resistant	5.5
SB2-134	C	Resistant	5.5
SB2-158	C	Resistant	5
SB2-161	C	Resistant	5.7
SB2-174	C	Resistant	4.5
SB2-259	C	Resistant	5.7
SB2-206	C	Resistant	6
SB2-225	C	Resistant	6
SB2-272	C	Resistant	6

I



Susceptible
SB2-276

Resistant
SB2-03

J

3K Rice genome project			
Type	Allele		
	T	C	T/C
Tropical Japonica	471	5	0
Temperate Japonica	360	0	3
Indica	18	1742	22
Aromatic	68	4	1
AUS	55	145	1
Admix	35	38	25

